Re-designing an anti-peptide antibody

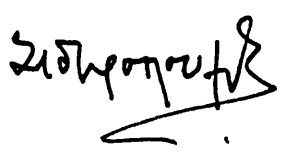
Sidiropoulou, Xenia

Award date:
1999

Awarding institution:
University of Bath

Re-Designing an Anti-Peptide Antibody

Submitted by Xenia Sidiropoulou
for the degree of
Doctor of Philosophy
Department of Biochemistry
of the University of Bath
1999



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

In the *de novo* design of antibody specificity the aim is to define a set of complementarity determining region (CDR) residues which is able to bind a pre-defined antigen. The starting point for this project was a monoclonal antibody (Gloop2) raised against residues 57-84 (the loop peptide) of hen-egg lysozyme (HEL) and the new target antigen was enkephalin. Both Gloop2 and enkephalin are structurally well characterised molecules through x-ray, and x-ray and NMR studies respectively.

Using a theoretical design procedure a docked complex between an all-alanine version of the Gloop2 combining site and enkephalin was constructed and, after reconstruction of side chains, twelve contact residues (7 in the heavy and 5 in the light chain) were identified as candidate positions for mutations. Each position was mutated to residue types, selected on the basis of both chemical redundancy and allowing for the fact that certain residues occur with high frequency in antibody combining sites. The resulting library was computationally screened using a simple force field and ten low energy complexes were selected.

The ten candidate antibodies were expressed as scFv fragments on the surface of filamentous phage creating a mini-library. The library was screened for binding to biotinylated enkephalin. After three rounds of selection, binding was observed and, DNA from the eluted binding phage was prepared and sequenced. It was found that the sequence matched one of the predicted mutation set. The scFv DNA of this selected design (GlaMor; Gloop2 originated antibody with specificity for the Morphine family molecules) was subcloned as a cassette from the phage to an expression vector. Different expression vectors were tested for efficient expression of soluble scFv in the periplasmic extract of *E. coli* cells. Gloop2 and GlaMor scFv's were purified from the periplasmic extract using immobilised metal affinity chromatography. The binding properties and specificity of Gloop2 against the loop peptide and enkephalin were studied using the surface plasmon resonance technology. However, severe problems in the expression of Glamor in any system explored prevented its full characterisation. The numerous attempts to resolve this expression problem are described in some detail.

Acknowledgements

This is the place to thank all the people that have made this thesis a reality.

I would like to thank my supervisor Professor A.R.Rees for his advice and support.
My temporary supervisor, during the last year, Dr. J.P.L.Cox for helping with practical matters, contacting people and obtaining useful tools for the work. Special thanks to Dr. J.S.Milner for spending time and getting me off to a good start with lab skills and techniques. Also, Dr .J.P.Petersen for establishing all the computational basis of the project and Dr. G.E.Elliott for starting the experimental work on it and constructing the phage library before I got involved. And also Nick in the graphics room for helping enormously in generating some of the modelling figures.

I would also like to thank Professor F. Rueker and Dr. R. Kerschbaumer, Institute of Applied Microbiology, University of Agricultural Sciences, Vienna, for their collaboration and advice.

Thanks to all the people of the Rees group, Lee for an excellent interaction, Muserref for always finding time to help, Rachel for constructive chat, and Herald for his sense of humour and also the Cox group, Claire, Simon and Mark for all their help, support and

warm atmosphere in the lab. Also, Alison Jones, Sue Phillips, Robert Griest and Marianne Rees offering me their scientific advice during all my lab years.

Other people in the department, Manos for his patience in "educating" me in computer skills and Jane for being always helpful with urgent bureaucratic matters. The University of Bath, for the Natural Sciences Demonstrator grant.

Also, people outside Bath, Dr. Walis Jones in BIAcore for helping me enormously with the binding assays and Professor R. Hawkins in Bristol University for providing the pUC119His6mycXba vector and the 9E10 antibody.

Finally I thank my parents and my sister in Athens, for their support.

Abbreviations

A	adenine or adenosine
AMP	adenosine monophosphate
AMPS	ammonium peroxisulphate
AP	alkaline phospatase
ATP	adenosine triphosphate
BIA	biomolecular interaction analysis
C	cytosine or cytidine
CDR	complementarity determining region
C_H	constant region of heavy chain
C_L	constant region of light chain
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid (disodium salt)
EDC	1-(3-dimethyaminopropy)-3-ethylcarbodiimide hydrocloride
Fab	fragment antibody
Fc	fragment crystallisable
FPLC	fats protein liquid chromatography
FW	framework
G	guanine or guanosine
gl2	Gloop2
HBS	HEPES buffered saline
HPLC	high performance liquid chromatography
Ig	immunoglobulin

IMAC	immobilised metal affinity chromatography
IPTG	β-D-isopropyl-thiogalactoside
kb	kilobase(s); kilobase pairs
K_D	binding constant
K_{on}	kinetic rate constant-association
K_{off}	kinetic rate constant-dissociation
NHS	N'-hydroxysuccinimide
NMR	nuclear magnetic resonance
OD	optical density
pI	isoelectric point
MCS	multiple cloning site
MW	molecular weight
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
RU	resonance unit(s)
RBS	ribosomal binding site
r.m.s	root mean square
SDS	sodium dodecyl sulphate
scFv	single chain Fv
T	thymine or thymidine
TBE	tris borate EDTA
TE	tris-EDTA buffer
TEMED	N,N,N',N'-tetramethyl-ethylaminediamine
u	unit(s)
V_H	variable region of the heavy chain
V_L	variable region of the light chain
(v/v)	volume: volume ratio
(w/v)	weight: volume ratio

Contents

8. Kinetic Analysis of Recombinant Antibody-Antigen Interactions

List of figures

List of tables

1

Introduction

1.1 The Immune System

Most infections in normal individuals are of limited duration and leave little permanent damage due to the individual's immune system that combats infectious agents. The immune system has two defence mechanisms:

a. natural (or innate) mechanisms that are present prior to exposure to infectious microbes or other foreign macromolecules and act as a first line of defense and,
b. adaptive (or specific) mechanisms that are activated when the innate response is not sufficient.

The term antigen is widely used to indicate any molecule, which can be specifically recognised by the adaptive elements of the immune system. The specificity of the immune response is due to lymphocytes, which are the only cells in the body capable of recognising and distinguishing different antigenic determinants. Lymphocytes

consist of distinct subsets of cells that are quite different in their functions and protein products, even though they all appear morphologically similar.

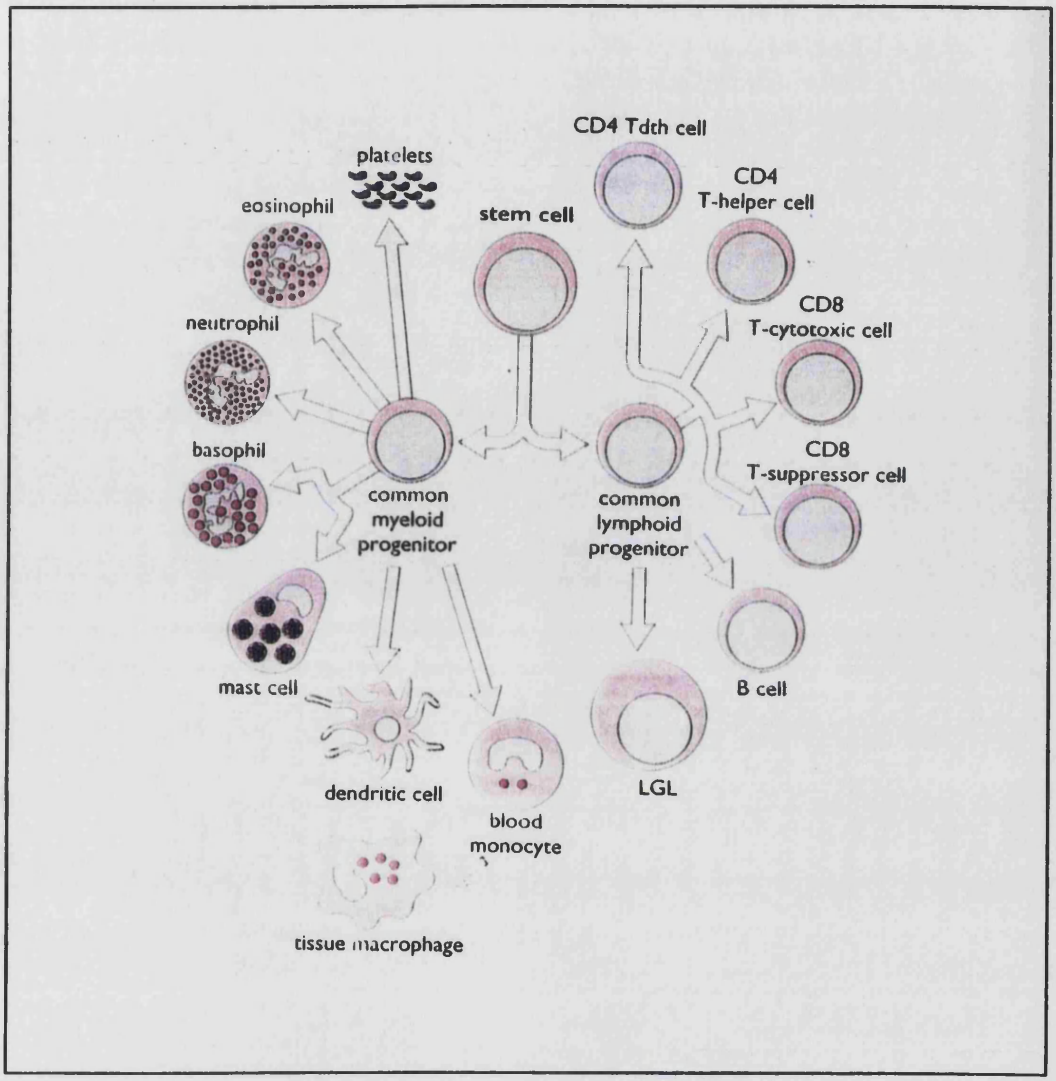


Fig. 1.1: The cells of the immune system
Two major categories of cells are derived from the common stem cell precursor: myeloid and lymphoid. Stem cells that differentiate to generate B cells reside in the bone marrow, and those that produce T cells reside in the thymus (From "Introducing Immunology", Mosby Edition)

T cells and B cells constitute the two important sets of lymphocytes that respond to antigens. B-lymphocytes express membrane bound and soluble antibodies while T lymphocytes express only membrane bound T cell receptor. The antigenic receptors carried, as membrane molecules on T lymphocytes are distinct but structurally related to antibodies, and bind to cell surface associated, but not soluble, antigens.

Antibodies are the major protecting agents of higher eukaryotes against foreign pathogens. A large repertoire of antibodies produced by B-cells provides recognition of almost all antigens and orchestrates their removal from the blood stream with the participation of the complement system and other immune system cells (T-cells, macrophages, phagocytes) (Fig. 1.1).

In the immune system the diverse repertoire of antibody molecules is generated by gene rearrangement, and is expressed (displayed) on the surface of B-cells, with each cell displaying a single-antibody species. Single B-cells proliferate on binding of antigen and differentiate to short-lived cells (plasma cells) making soluble antibody or to long-lived cells (memory cells) with surface antibody (Fig. 1.2).

First, the antigen encounters the pool of B cells, comprising of memory cells and *naive cells*, expressing a continuously changing portion of the potential germline gene repertoire. *Primary responses* (affinities 10^5-10^7 M^{-1}) start largely on the naive portion of the repertoire, which is a much more diverse repertoire, whereas secondary responses start with the cells of the memory pool. *Somatic hypermutation* provides further variation (10- to 200- fold), from which antigen selects proliferative cells with improved binding characteristics. As immunisation proceeds, antibodies arise that

come quite late in the immunization and which do not seem to be related to the pool of the initially reacting ones. It can be that these are very uncommon components of the naive pool, which eventually appear by pure chance, and having such high affinity for antigen, are able to survive and compete with already mature high affinity cells from the memory pool.

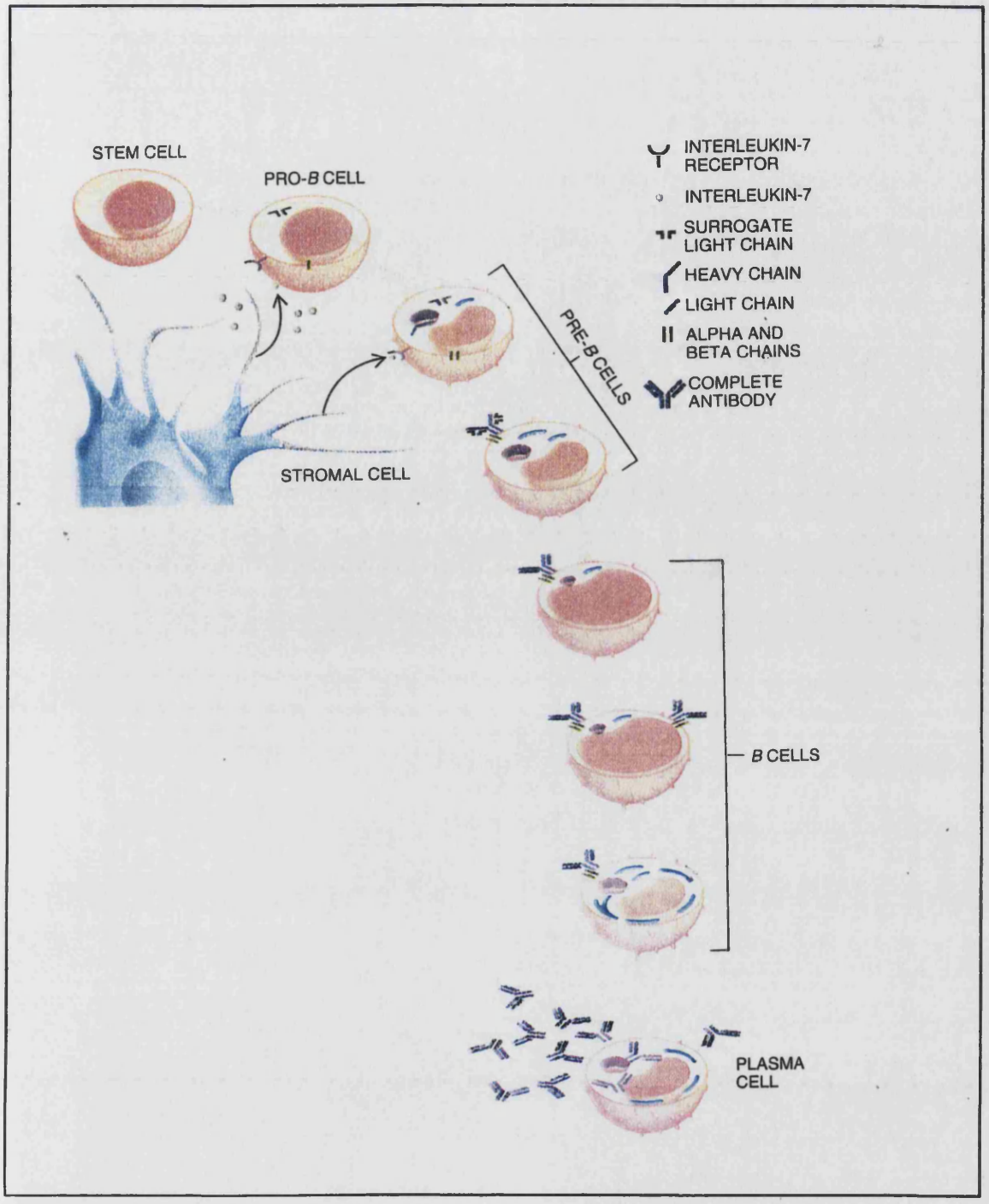


Figure 1.2: The development of B cell
It starts with stem cell in the bone marrow. These cells reproduce themselves and also spawn lineage that pass through pro-B and pre-B stages to become B cells. B cell development culminates in the plasma cell, which secretes antibodies or to long-live memory cells (not shown), which retain their surface Ig and are involved in future immune response to the same or cross reactive antigen (From Scientific American, September 1993)

1.2 Antibodies

1.2.1 Antibody Structure

The immunoglobulin G (IgG) molecule consists of a tetramer of two identical 25kDa polypeptides (the light chain, L) and two identical 50kDa polypeptides (the heavy chain, H). Crystal structures of IgG and IgG fragments have shown the antibody combining site (ACS) to be formed by the juxtaposition of six hypervariable loops or complementarity determining regions (CDRs), three from the light chain variable (VL) and three from the heavy chain variable (VH) domain. The CDRs of each chain are supported on a framework region (FR1, FR2, FR3, and FR4) which consists of conserved β–strands that fold to form a β–sandwich. When a light and heavy chain come together, one surface of each sandwich associates to form a β–barrel structure. Supported on this β–barrel scaffold the six CDRs, pack together in a tertiary structure to form a relatively flat platform with a surface area of about 700Å (Fig.1.3). The variable domains (VH and VL) contain all the determinants of antigen recognition they represent two of six domains in the IgG molecule, the remaining four being Constant or C-type domains (CH1-CH3 and CL). The relative positions of these domains form the characteristic Y-shaped IgG molecule (Fig. 1.4). The smallest antigen-binding fragment is the Fv (fragment variable), which consists of the VH and VL domains while the constant domains comprise the Fc (fragments crystallisable) fragment that has effector functions. The Fab (fragments antigen binding) fragment consists of the VH-CH1 and VL-CL domains covalently linked by a disulphide bond and more recently, the scFv (single chain Fv) antibody fragment (the VH and VL domains linked by a flexible peptide chain) was developed to overcome the tendency

of the non-covalently linked VH and VL domains to dissociate during expression studies (Bird *et al.*, 1988 and Huston *et al.*, 1988). The polypeptide can link either the C-terminus of the VH to the N-terminus of the VL or the C-terminus of the VL to the N-terminus of the VH. As long as the linker is of adequate length, a wide range of different linker sequences can be tolerated (Huston *et al.*, 1991). scFvs typically have affinities similar to the antibody from which they are derived (Bird and Walker, 1991, Fig. 1.5).

Porter (1958) and Edelman (1970) established the four-chain structure of antibodies for an IgG in the early 1960s. The three-dimensional arrangement of the heavy and light chain sequences as constant and variable domains was established for the Fab fragment in 1973 by Poljak *et al.* and for the Fc by Diesenhofer *et al.* 1976.

The six CDR loops forming the immunoglobulin binding site have different sequences in different antibodies; the remainder of the V domain exhibits less sequence variation and forms the framework regions (Fig. 1.6). The affinity and the specificity of a binding site is determined by the residues in both the hypervariable regions and the adjacent framework regions (Jones *et al.*, 1986). The main chain conformations of the six hypervariable regions are determined to some extent by loop length (de la Paz *et al.*, 1986) and by the interactions of a few residues at specific sites in the hypervariable regions and for certain loops in the framework regions. Hypervariable regions, that have the same conformations in different immunoglobulins have the same or very similar residues at these sites (Chothia *et al.*, 1989). An examination of the library of the known immunoglobulin sequences showed that many immunoglobulins have hypervariable regions that are the same size as those in the

known structures and contain the same or closely repeated residues at the sites responsible for the known conformations (Chothia and Lesk, 1987). These observations indicated that for at least five of the hypervariable regions there is a small repertoire of canonical main-chain conformations and that the conformation actually present can be predicted from the sequence by the presence of specific residues.

The antigen-binding activity can be grafted from one antibody to another by transplanting the CDRs (Jones *et al.*1986, Verhoyen *et al.* 1988), showing that the β–sheet is a structural framework for the CDRs, that the CDRs are responsible for binding to antigen, and that the same β–sheet framework can provide a scaffold for mounting different antigen loops. CDR grafting has been used to humanise rodent antibodies for human therapy (Riechman *et al.*, 1988).

1.2.2 Antibody diversity and Gene Organisation

The domain structure of the antibodies is mirrored at the level of the gene, as the individual domains are encoded by separate exons. As we know today the variable region of antibodies (and similarly for T-cell receptors) is the result of the combinatorial recombination of three families of gene fragments (V, D, and J). The first and second hypervariable loops (H1, H2 and L1, L2) are encoded by the germ-line V genes, but the third hypervariable loop by the combination of V, D and J elements (H3 for the VH domain) and V and J elements (L3 for the VL domain). The third hypervariable loop of VH is accordingly the most diverse and is often the longest of the loops. The joining DNA fragment is not precise, and this gives rise to variations in length, particularly around the D segment which explains why the H3 hypervariable region is so variable. L2 is often not used in binding to smaller antigens even when all the other loops are used for binding (Tormo *et al.*, 1994)

The potential diversity of the mouse primary repertoire is huge ($>10^{10}$) as a consequence of the combinatorial arrangements of the genetic elements with only a fraction of the potential repertoire available through the limited number of clones (10^7-10^8) expressing antibodies. In the secondary response genes used from the primary pool undergo hypermutation followed by selection and resulting in higher affinities antibodies. The rate of mutation approaches 10^{-3} or $3x10^{-4}$ mutations per base pair per cell division and is localised to the segment coding for the variable portion of the antibody genes. As immunisation proceeds, additional high-affinity antibodies

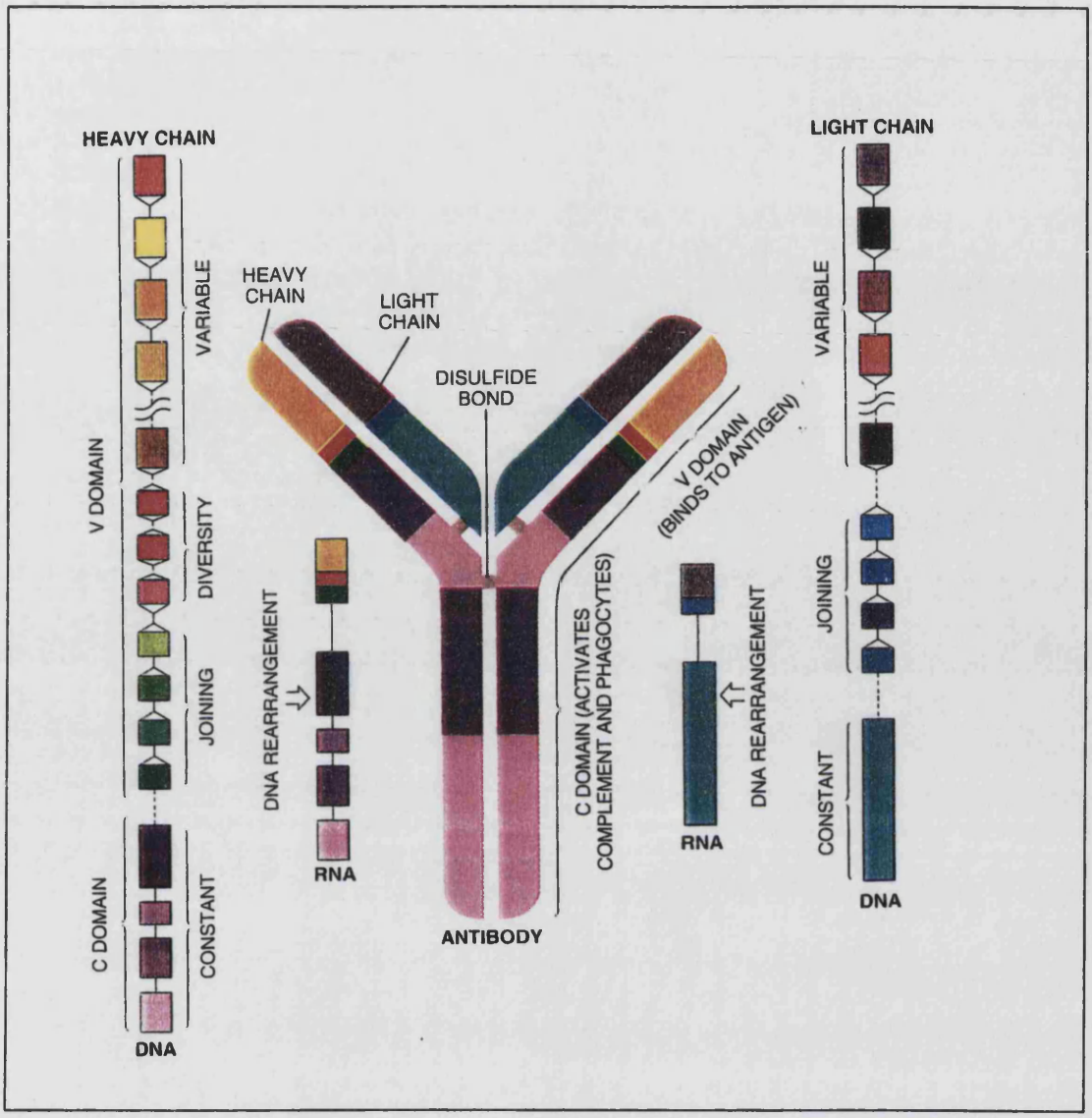


Fig. 1.4: The antibody molecule
Genes that consist of different DNA segments encode the chains. These segments rearrange to make genes for chains that are different in each B cell. The joining is so variable, so that only a few gene segments generate the estimated 100 million distinct antibodies the body is capable of producing (From Scientific American, September 1993).

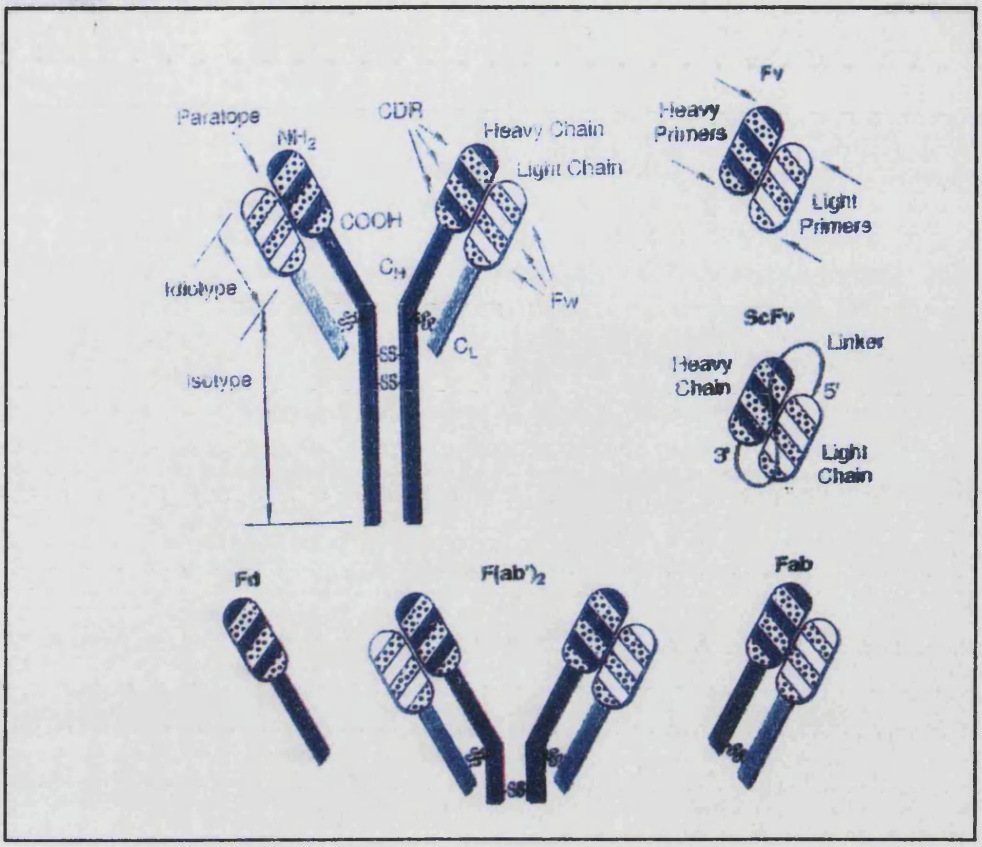


Figure 1.5: Antibody fragments
Antibody model showing the subunit composition and domain distribution along the polypeptide chains. Fragments generated by the proteolytic cleavage and/or recombinant technology appear on the right and bottom part of the figure (From Pharmacia Handbook).

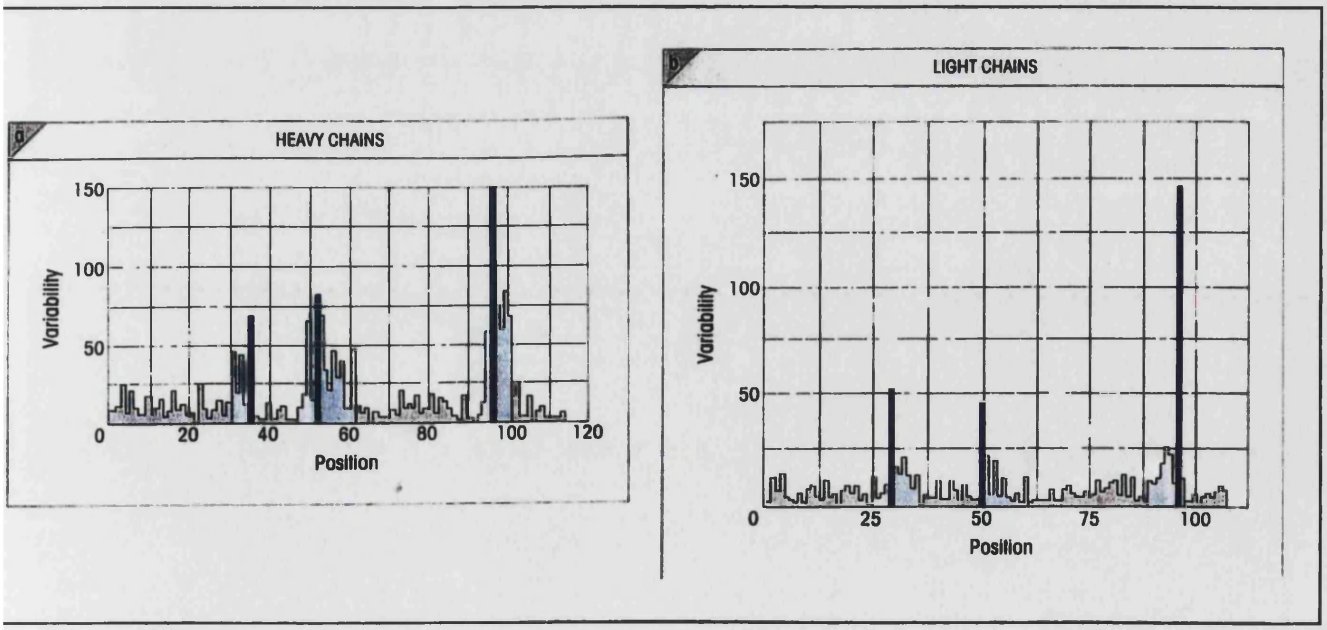


Fig. 1.6: Plot of amino acid variability in the variable region of immunoglobulin heavy and light chains.
The sequences from a large number of myeloma monoclonal proteins are compared and variability at each position is computed as the number of different amino acids found divided by the frequency of the most common amino acid; the higher the number the greater the variability. The three CDRs in the heavy (a) and light (b) chains are shown in blue, while the intervening peptide sequences are termed framework regions (in gray).

1.2.2 Antibody diversity and Gene Organisation

The domain structure of the antibodies is mirrored at the level of the gene, as the individual domains are encoded by separate exons. As we know today the variable region of antibodies (and similarly for T-cell receptors) is the result of the combinatorial recombination of three families of gene fragments (V, D, and J). The first and second hypervariable loops (H1, H2 and L1, L2) are encoded by the germ-line V genes, but the third hypervariable loop by the combination of V, D and J elements (H3 for the VH domain) and V and J elements (L3 for the VL domain). The third hypervariable loop of VH is accordingly the most diverse and is often the longest of the loops. The joining DNA fragment is not precise, and this gives rise to variations in length, particularly around the D segment which explains why the H3 hypervariable region is so variable. L2 is often not used in binding to smaller antigens even when all the other loops are used for binding (Tormo *et al.*, 1994)

The potential diversity of the mouse primary repertoire is huge ($>10^{10}$) as a consequence of the combinatorial arrangements of the genetic elements with only a fraction of the potential repertoire available through the limited number of clones (10^7-10^8) expressing antibodies. In the secondary response genes used from the primary pool undergo hypermutation followed by selection and resulting in higher affinities antibodies. The rate of mutation approaches 10^{-3} or $3x10^{-4}$ mutations per base pair per cell division and is localised to the segment coding for the variable portion of the antibody genes. As immunisation proceeds, additional high-affinity antibodies

gradually emerge with V-D-J combinations, which are rarely found in the primary repertoire (repertoire shift).

1.2.3 Combining site topography and CDR conformation

The CDR residues number only around 70 of the total approximately 230 residues of the Fv antibody fragment and for some CDRs adopt only a limited number of canonical backbone conformations (Chothia *et al.*, 1989), determined by their loop length (de la Paz *et al.*, 1986) and certain key, "structurally determining" residues **(canonical residues)**. Chothia (1991) suggested that there is only a small repertoire of main-chain conformations (canonical structures) for at least five of the six hypervariable regions, that sequence variations within the hypervariable regions modulate the surface that these canonical structures present to the antigen and, that sequence variations within the hypervariable and the framework regions shift the canonical structures relative to each other by small but significant amounts. The adaptability of the antibody framework has been well illustrated, as variations in loop size, amino acid composition and relative domain disposition can tailor the antibody binding site to obtain the specificity and affinity in the immune recognition process (Wilson and Stanfield, 1993). The essential validity of the canonical concept has been demonstrated by the reasonable success with which it has been used to predict the structure of the antibody binding site, for up to 5 of the 6 CDRs, before the experimental determination (Chothia *et al.*, 1989).

Webster *et al.* (1994) suggested three types of combining site: cavity (hapten), groove (peptide, DNA, carbohydrate) and planar (protein) and this topographic classification

was based on the analysis of 20 X-ray structures. The amount of Fab surface buried on complex formation increases in the order cavity<groove<planar while the percentage of antigen surface area buried in the interface area buried decreases as the size of the antigen increases.

There is no totally satisfactory way of predicting which CDR combinations give rise to which surface topography. Analysis of antigen contacting residues for ten pairs of complexed and uncomplexed antibody-antigen crystal structures clustered the antibody combining site surfaces into four topographic classes: concave and moderately concave (mostly hapten binders), ridged (mostly peptide binders) and planar (mostly protein binders). However, when the whole antibody combining site was subjected to the same analysis the trend was much less marked, due to an increase in the noise resulting from the inclusion of non-accessible (or non contacting residues) of the non-interface surfaces in the calculations (MacCallum *et al.*, 1996). The non-contacting residues within the CDRs coincided with residues identified by Chothia *et al.* (1989) in defining canonical backbone conformations.

1.2.4 Antibody-Antigen Interactions

The Thermodynamic View

In general two macromolecules in solution have to overcome large entropic barriers before they can form a tight association. There is the loss of the entropy of free rotation and translation of the separate molecules, and there is the loss of conformational entropy of mobile segments and of side chains upon binding. On the other hand, entropy is gained when water molecules are displaced from the surfaces that become the new interface.

Enthalpic contributions arise from van der Waals interactions together with the more specific hydrogen bonds and salt bridges. As one would expect, the larger the ligand the more surface area is involved on both molecules, although the surface area does not increase proportionally to the total surface area of the ligand (Davies and Padlan, 1990).

The Role of Water

The effect of water displacement from the antibody-antigen interface and the entropy gain is quite significant. In the structures observed by X-ray analysis, at 2.5Å resolution water molecules have been identified at the interfaces (Tulip *et al.*, 1992) of Fab complexes unfortunately with limitations in the certainty with which ordered molecules could be located. In other studies at 1.8Å resolution it was demonstrated that when comparing structures of free and bound Fv fragment of D1.3 several of the water molecules in the free antibody combining site were retained and that additional

water molecules linked antigen and antibody upon complex formation Bhat *et al.* (1994).

It has also been suggested that water molecules present at the interface, may mediate antibody-antigen interactions by increasing the packing density and contributing to charge complementarity (Tulip *et al.*, 1992). Water molecules have been found in the interface of the anti-lysozyme antibody D1.3 (Mariuzza and Poljak, 1993), around the periphery of some anti-protein interfaces (Tulip *et al.*, 1992) and in the combining site of some anti-peptide antibodies (Shoham, 1993). It is a reasonable speculation that, where complementarity is imperfect, filling-in by water molecules might occur. Recently, Corell and Mallqvist (1997), calculated that interfacial crystal water molecules contributed 25% of the total calculated binding strength when assessing the origins of binding strength and specificity for three crystal determinants. They calculated changes in binding free energy (calculated value) and the experimental values (observed) for single and double mutations of these complexes. They proposed a model where the observed differences in the binding strength (between the experimental and the calculated values) were attributed to indirect changes due to released crystallographic water molecules that were near the mutation sites.

The antigen-binding site

The specificity of the antibody-antigen interaction might be regarded as a paradigm of molecular recognition in general and is believed to result from the complementarity of the interaction surfaces. The investigations reviewed here confirm that there is remarkable shape complementarity; depressions on one surface are filled by protuberances from the other, often leaving no holes large enough for even water

molecules. Hydrogen bonding is another factor that makes the interaction specific. Of particular relevance here is the directionality of the hydrogen bonds, necessitating a hydrogen bond receptor within a certain distance and within a certain solid angle of the hydrogen bond donor in order to form a strong bond.

When the antibody-antigen interfaces are inspected, a high degree of shape complementarity is often seen, particularly for protein complexes. In an analysis of the interface of the influenza virus N9 neuraminidase-NC41 Fab complex it was found that the average packing density of atoms at the interface was lower than for a typical protein core (Tulip *et al.*, 1992). However, a second analysis by Walls and Sternberg (1992) found that the packing densities of the three anti-hen egg-lysozyme antibody interfaces were as high as the protein cores. In contrast, electrostatic complementarity is not always as high (Novotny and Sharp, 1992).

Antibody-antigen interactions

Affinity constants (K_D or K_A) and kinetic binding constants (k_{on} and k_{off}) are used to quantitate the antibody-antigen (Ab·Ag) interaction. Two approaches are typically used to measure affinity constants.:

- the determination at equilibrium of the ratio of the concentrations of the antibody (or antigen) free and engaged in complex and
- measurement of the kinetic dissociation and association constants, after which K_d can be obtained from the ratio k_{off}/k_{on}.

Amino acid substitutions in antibody-antigen interfaces play an important role in affinity maturation of antibody responses and in antigenic variation. These changes

have the capacity to drive the affinity towards more tightly bound complexes or can effectively abolish the interaction entirely, providing a mechanism for antigenic variation. In studies on the effects of point changes in the binding for the influenza virus neuraminidase (Tulip *et al.*, 1992) shown that two separate mutations located near the edge of the combining site had the same effect on the antibody-antigen interface as the amino acid substitutions within the interior of protein molecules (Matthews, 1991). Folding could be considered to be more tolerant of amino acid substitutions because of the cooperative effects caused by the interactants being convalently linked to each other.

In other studies Padlan (1990) showed that Tyr residues are unusually common in the regions of antibodies responsible for contact with antigens. Tsumoto *et al.* (1995) clarified this by studying the interaction between hen egg lysozyme (HEL) and its monoclonal antibody HyHEL10, both having well-characterised structures. Four Tyr residues in the VH chain were replaced by Ala, Leu, Phe, or Trp making in total 16 mutated constructs. Twelve of them could be expressed, but two mutants could not be obtained in *E. coli* expression system and a further two mutants could not be purified by affinity chromatography. The Tyr residues at each mutated site made positive contributions to the interaction to different degrees.

Studies on antibodies against various haptens concluded that residue 96L interacts directly with the antigen (Roberts *et al.*, 1994, Riechmann and Weill, 1993) while an Arg at position 96 correlates with ds DNA binding by autoantibodies to DNA (Marion *et al.*, 1992).

Conformational Changes

One of the more controversial issues in the antibody field has been the question of whether conformational changes occur in the antibody or antigen upon binding (Webster *et al.*, 1994). This issue has been touted as a lock-and-key versus an induced fit contest. Two Fab structures, an anti-DNA (Herron *et al.*, 1991) and an anti-peptide (Rini *et al.*, 1992), illustrated the extensive nature of the conformational changes that can occur in an antibody following the 'induced fit' theory. These might range from small changes in side chain orientation (Stanfield *et al.*, 1990), to segmental changes in the main chain (Stanfield *et al.*, 1990), to CDR loop rearrangements (Herron *et al.*, 1991), to the more extensive alterations in the relative disposition of the VL and VH domains which form the antibody combining site (Herron *et al.*, 1991).

1.3 Antibody Engineering

1.3.1 Production of Polyclonal and Monoclonal Antibodies

The blood of an immunised individual contains many different antibodies, each derived from a particular clone of B cells and each having a distinct structure and specificity for antigen. Nevertheless, during early immunological experiments antibodies were sufficiently similar to each other that it was possible to purify mixtures of antibodies away from other blood proteins. Working with these mixtures, immunologists were able to deduce the overall structure of antibody molecules. However, the molecular heterogeneity of these **polyclonal antibodies** interfered with more detailed analysis of antibody structure, such as amino acid sequence determination. The key methodological breakthrough in this endeavour was the discovery that patients or animals with multiple myeloma, a monoclonal tumour of antibody-secreting plasma cells, often have high levels of biochemically identical antibodies or portions of antibodies in their blood or urine. These immortalised cells provide a source of individual antibody-secreting cells from an immunised animal, permitting the selection of individual **monoclonal antibodies** of predetermined specificity.

The monoclonal antibody method subsequently developed involved cell fusion or somatic cell hybridisation between a normal antibody-producing B cell and a myeloma line using a fusogenic agent such as polyethylene glycol. If the conditions are carefully controlled, some cells fuse and some of these both make specific

antibody and will survive indefinitely in culture. The unfused myeloma cells eventually die under the genetic selection conditions used to favour the fused progeny, while the unfused primary B-cells die within a few days. Eventually hybridoma cells can be grown individually and the clones used to establish permanent large cultures from which the monoclonal antibodies can be harvested (Kohler and Milstein, 1975)

1.3.2 Antibody Display Technology

The polymerase chain reaction, described in 1988 (Suggs *et al.*, 1988), greatly simplified the task of obtaining and cloning immunoglobulin DNA and has made it possible to reproduce antibody gene repertoires present *in vivo*. Design of primers for the 3' end of immunoglobulin variable region genes was straightforward since primers could be based on the constant regions, all of which have been sequenced. Design of primers for the 5' end of the V gene was less straightforward due to the sequence variability of different V-genes, although it was known that primers did not have to be a perfect match for the template DNA. In the earliest attempt to use PCR to amplify V-genes, N terminal protein sequencing was done on purified antibody from a hybridoma and the sequence used to assign the VH and VL gene families. The VH and VL gene assignments were used to design degenerate primers based on FR1.

A more generally applicable approach was taken by Orlandi *et al.* (1989) where the nucleotide sequences of murine VH and VL genes were extracted from the Kabat database (Kabat *et al.*, 1987) aligned and the frequencies of the most common nucleotide regions plotted for each position. Conserved regions were identified at the 5' and 3' regions of the VH and VL genes, and the sequences of these regions were

used to design oligonucleotide primers. Restriction sites were incorporated into these primers to permit cloning directly into vectors for sequencing and expression in eukaryotic cells. VH and VL genes amplified from a hybridoma were expressed in eukaryotic cells, and functional recombinant antibody was produced, verifying this method (Orlandi *et al.*, 1989). PCR with immunoglobulin specific primers could also be used to amplify genomic VH and VL gene repertoires from DNA or RNA prepared from mouse spleens (Huse *et al.*, 1989, Gussow *et al.*, 1989, Ward *et al.*, 1989). Sequence analysis of the amplified VH and VL genes indicated that a diverse gene repertoire could be produced (Gussow *et al.*, 1989).

When these gene libraries are displayed on filamentous bacteriophages, antibodies can be made completely *in vitro*, by-passing the immune system and the immunisation procedure, and allowing *in vitro* tailoring of the affinity and specificity of the antibody. VH and VL genes can be fused to the amino terminus of the phage minor coat protein pIII. The technology has been used to display both scFv and Fab antibody fragments. McCafferty *et al.* (1990) demonstrated the feasibility of displaying fragments of antibodies on the surface of a bacteriophage and that these fragments folded correctly and bound antigen. Antibody variable-region genes were cloned into the gene encoding the minor coat protein (gene3) of the filamentous bacteriophage fd (Fig. 1.7). The fusion protein created, consisting of the antibody fragment at the N-terminus of the coat protein is incorporated into the phage particle (a "phage antibody"). Each recombinant phage genome contains the DNA encoding the specific antibody displayed on its surface, allowing the phage particle carrying antibody gene to be selected directly using the binding properties of the expressed protein. This discovery opened a novel route for antibody isolation.

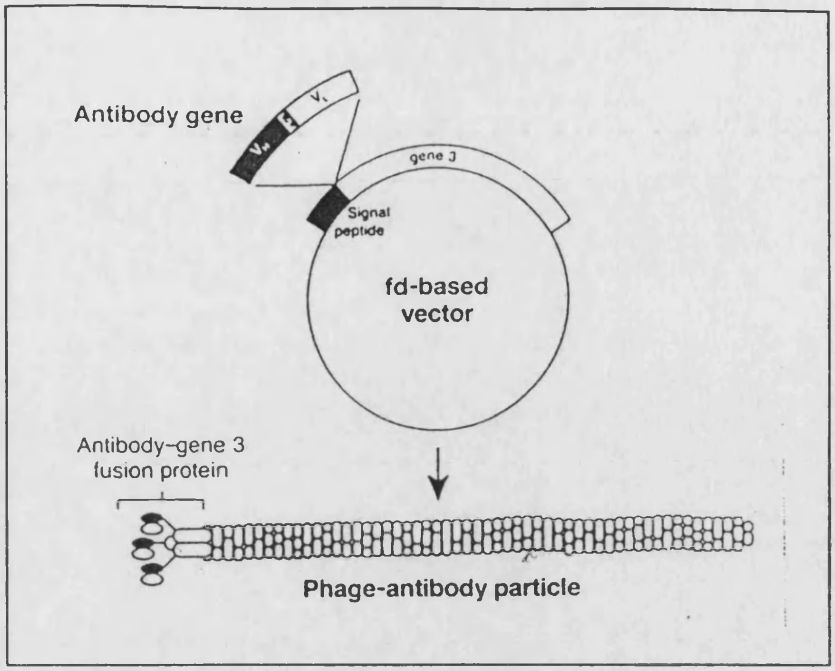


Fig. 1.7: Phage antibody particle
The filamentous bacteriophage (fd) contains a 6.4Kb single-stranded, circular DNA molecule associated with 2800 copies of the major coat protein encoded by gene 8. At the end of the phage particle there are three copies of the minor coat protein encoded by gene 3. Antibody-gene sequences have been inserted at 5' end of the gene 3 protein. This fusion protein is assembled into phage particles which display antibody fragments on the surface (Chiswell and McCafferty, 1992).

Repertoires of antibody genes are amplified using PCR and cloned into phage, thus creating a large library of phage each displaying a specific antibody. Within these large libraries of phage antibodies (potentially representing between 10^6 and 10^9 different antigen-binding specificities) each phage expresses an individual heavy-and light-chain combination (Fig 1.8), although these combinations do not necessarily reflect the *in vivo* pairings since VH/VL scrambling occurs during the library construction. The linkage between antibody genotype and phenotype allows the enrichment of antigen-specific phage antibodies, using immobilised or labelled antigen (e.g via antigen bound to a solid surface such as a Sepharose column or a coated tube). Phages that display a relevant antibody are retained on a surface coated with antigen, while non-adherent phage are washed away. Bound phages can be recovered from the surface, reinfected into bacteria and re-grown for further enrichment and, eventually, for binding analysis. The phage antibody can be both

analysed and used directly from the culture supernatant as a reagent in techniques such as ELISA (Fig 1.9).

The success of antibody phage display hinges on the combination of this display and enrichment method, with the creation of large combinatorial repertoires of phage antibodies. Each of these antibodies has a different combining site, generated by a PCR-based amplification of antibody V-genes (Winter *et al.*, 1994). An antigen-driven enrichment procedure can be used to isolate even the rarest phage antibody, for example where only one in 10^7 phage in a population is specific for the antigen. This basic technique has been applied to construct very large and highly diverse combinatorial repertoires of antibody fragments, from which antigen-specific antibodies to any chosen antigen can be selected (Griffiths *et al.*, 1994). It has also been possible to use antibody-gene VIII protein fusion in which, theoretically, over 2000 copies/phage particle could exist (in practice much less than this), leading to a multivalent selection system. This latter approach involves the display of greater numbers of antibodies on each phage particle as well as a variable number of antibodies on different phage particles produced by the same cell. Selection would therefore be a function of avidity (i.e. a combination of affinity and valency) rather than affinity alone, limiting the usefulness of this approach, particularly when selecting high affinity antibodies. Antibody fragments using a display system based on the gene VIII product have been engineered onto phages for Fab fragments by Kang *et al.*, 1991, Chang *et al.*, 1991 and Huse *et al.*, 1992)

Other groups have used gene3- display to isolate antibodies derived from libraries originally constructed on λ bacteriophage (Huse *et al.*, 1989). Plaques expressing

Other groups have used gene3- display to isolate antibodies derived from libraries originally constructed on λ bacteriophage (Huse *et al.*, 1989). Plaques expressing antigen binding antibody fragments were identified by obtaining plaque lifts that were then probed with labelled antigen. This type of antibody expression system has also been successfully applied to the isolation of murine antibodies specific for influenza virus hematoglutin (Caton and Koprowski, 1990) and for human Fab fragments specific for tetanus toxoid (Persson *et al.*, 1991).

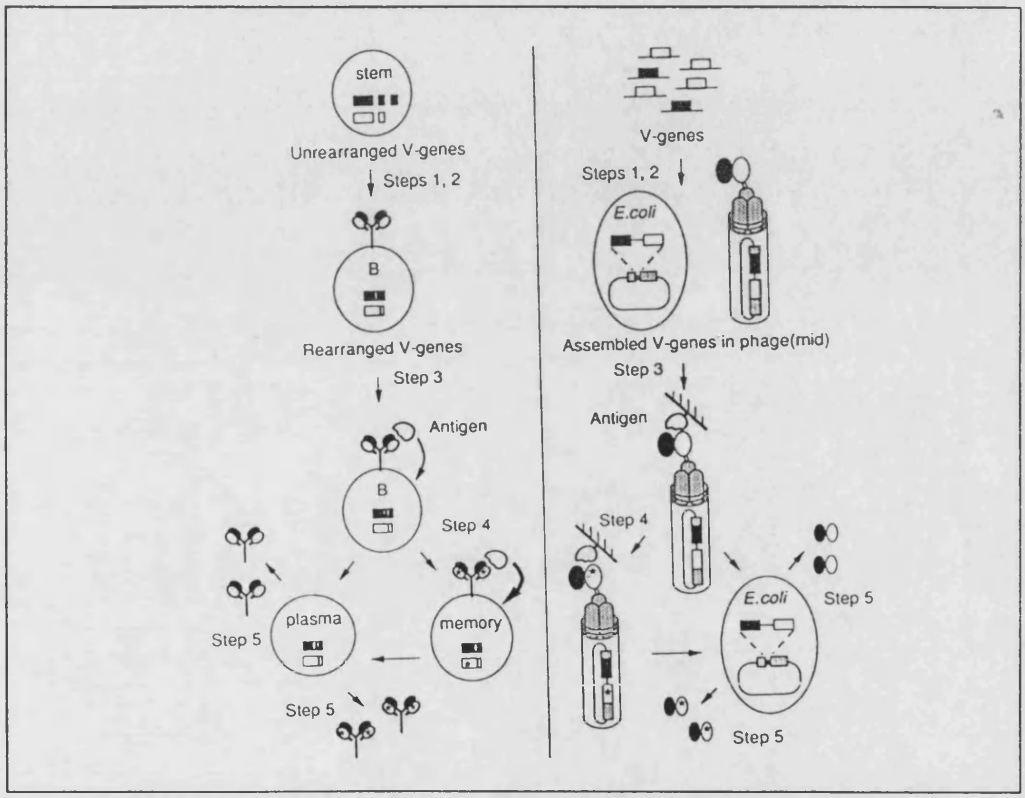


Fig. 1.8: The strategy of the immune system in vivo (left) and using phage (right)
Step 1: rearrangement or assembly of germ line V-genes; Step 2: surface display of antibody; Step 3: antigen-driven or affinity selection; Step 4: affinity maturation; Step 5: production of soluble antibody (or antibody fragment) (Hoogenboom *et al.*, 1992).

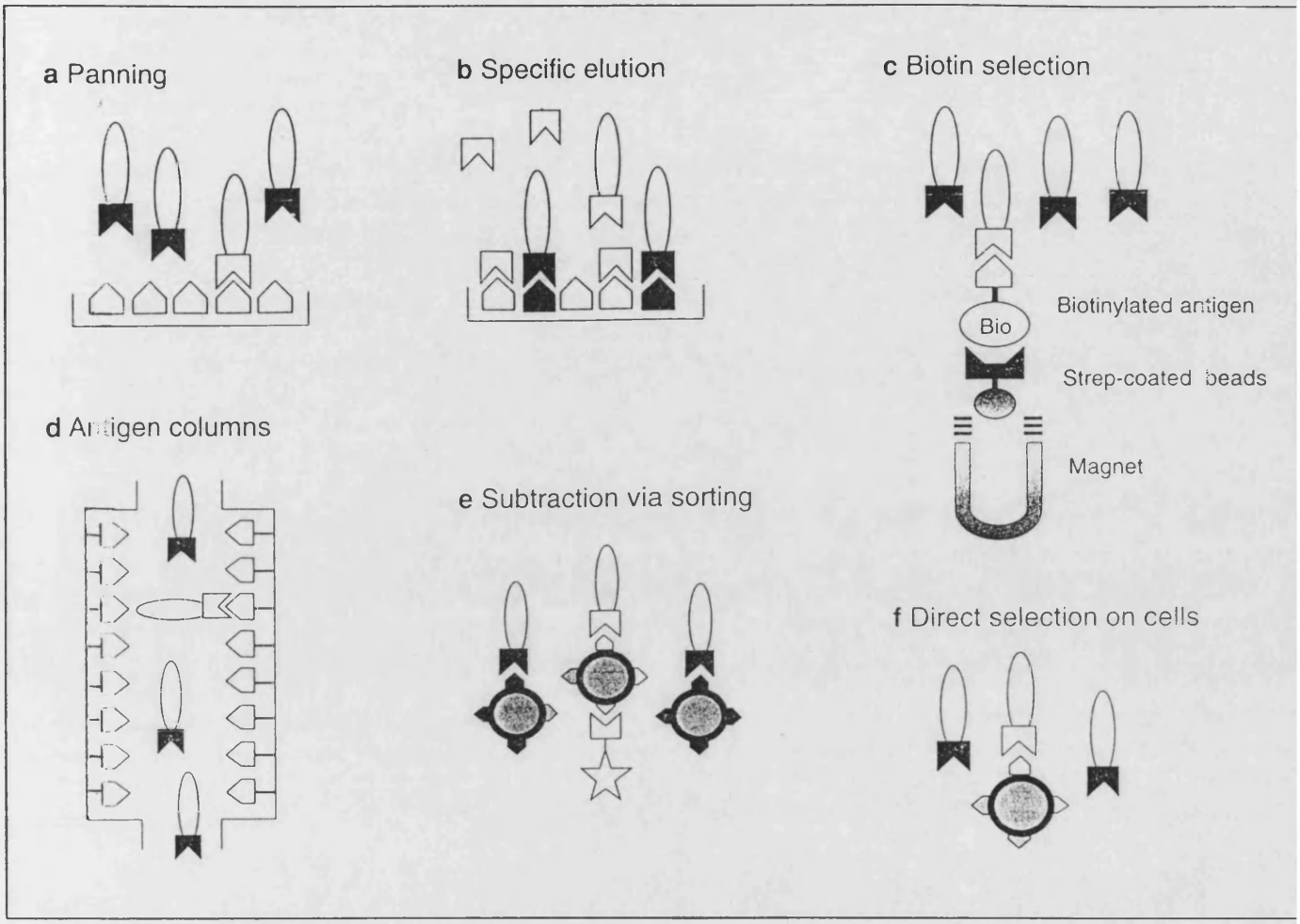


Fig. 1.9: Selection methods
Phages displaying antigen-specific antibodies (white) are usually separated from non-binding phages (black) using the above procedures (Hoogenboom, 1997)

A recent application of selectively infective phage (SIP) (Fig.1.10) has been devised, in which cognate pairs of peptidic antigen and antibody can be selected at the same time (Krebber *et al.*, 1995 and for a Review, Spada *et al.*, 1997). In the SIP system, the antibody-displaying phage is itself non-infectious and acquires the ability to infect cells only upon the occurrence of a cognate interaction between antibody and antigen in the periplasm, directly liberating infectious phage from the host.

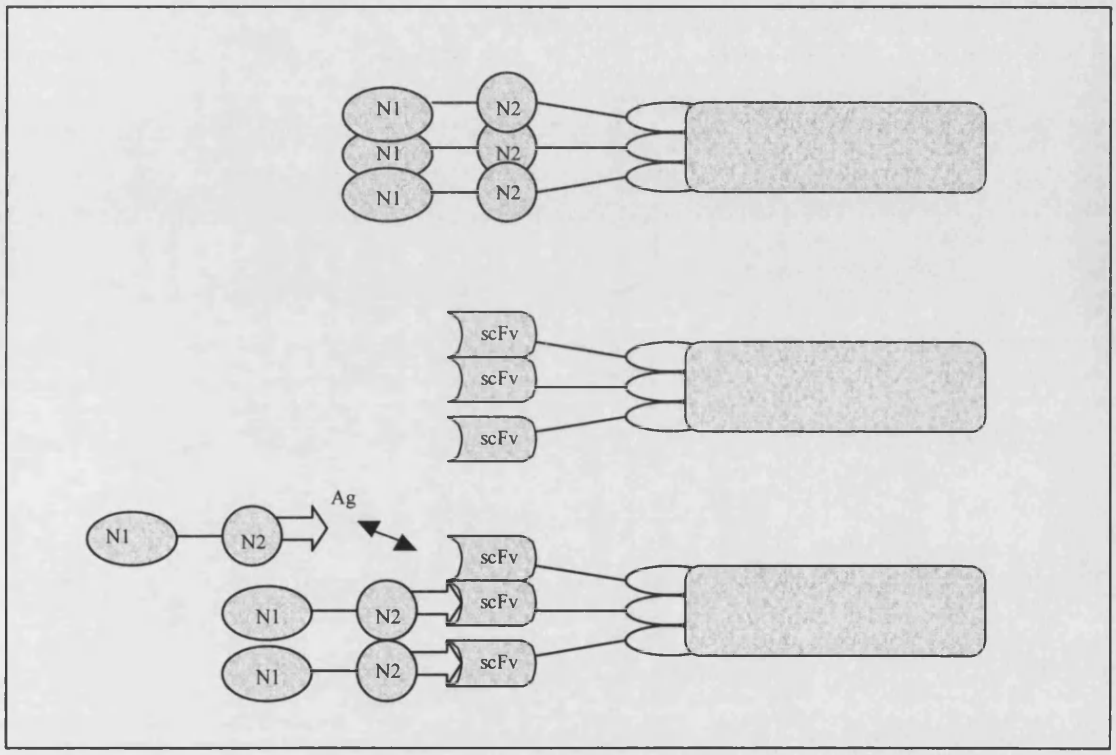


Fig.1.10: Selectively Infective Phage (SIP)
(Top) wt fd phage in infective
(Middle) Multivalent antibody-displaying phage, devoid of N1-N2, is non-infective
(Bottom) Infectivity is restored by the cognate interaction between antibody and antigen, reconstituting of gIIIp functions (Krebber *et al.*, 1995).

1.3.3 Expression of Recombinant Antibody Fragments.

Different expression systems have been studied for the production of recombinant antibodies. These range from bacteria, insect cells and plants to transgenic animals, with each system having its own advantages and problems. Different antibody fragments (Fab, Fv and scFv) have been expressed in all of these systems as well as whole chimeric antibodies.

E. coli offers an attractive system for the expression of antibody fragments and has major advantages over production in eukaryotic cells since, yields are higher, costs are lower and the system is more easily regulated and manipulated. However, in other studies the mammalian expression systems appear to be more successful for the production of correctly folded intact antibodies or antibody derivatives (Gilliland *et al.*, 1996).

Various fragments that retain antigen-binding activity can be created. An Fab construct most closely resembles the native binding site structure, retaining the adjacent constant region domain for both heavy and light chain. However, the variable region domains can also be expressed separately and still associate to form an Fv fragment. It has also been demonstrated that the variable regions can be joined by a short linker segment, thereby providing a single chain binding site or scFv.

For some biotechnological applications, scFv's are better suited than whole antibodies because they are smaller in size yet they retain their antigen binding specificity. For example, in tumour targeting studies they have been shown to exhibit increased

tumour penetration and localisation in comparison to intact mAb or Fab fragments (Colcher *et al.*, 1990). In addition, scFvs can be made "immunosilent" by CDR grafting (Verhoeyen *et al.*, 1988) or "resurfacing" (Roguska *et al.* 1994) that reduce the foreign nature of the proteins

Fab and Fv fragments have also been expressed in *E. coli* (Field *et al.*, 1988;1989) and the final choice of structure, depends on the specificity requirements and the proposed use. For example, an Fv fragment has been shown to have a half life of only 1.3 hours at 37°C rendering it unsuitable for most studies. The addition of a flexible linker segment (scFv) has been shown to increase the thermostability to 15 hours (Glockshuber *et al.*, 1990). This significant increase in stability is thought to be a function of limiting the dissociation of VH and VL. However, the constant regions in an Fab also limit dissociation, such that an indefinite half-life is observed at 37°C. Other structural modifications to increase the stability of smaller fragments, such as introduction of additional disulphide bonds, may prove useful for increasing the thermostability (Glockshuber *et al.*, 1990).

The affinity of the expressed antibody fragment can also vary with respect to structure. No difference in affinity has yet been observed between chemically cleaved Fab and genetically constructed, bacterially expressed Fab (Buchner *et al.*, 1991). Similarly Skerra and Plünkthun reported that their Fv directed against phosphorylcholine retained the same affinity as that of the parent antibody (Skerra and Plünkthun, 1988). In contrast the addition of a linker to one Fv to form a scFv was found to result in a 2-4-fold decrease in the apparent affinity (Glockshuber *et al.*, 1990). Other scFv's have also been shown to have affinities as low as one tenth of

their Fab counterparts (Huston *et al.*, 1988a). Such is not always the case and may be a function of the specific linker construct utilised hindering the proper formation of, or access to, a particular binding site (Bird and Walker, 1991). It would seem reasonable to expect that this may be less of a concern for binding sites directed toward small haptenic determinants. Interestingly, however there are examples of scFvs directed against small chemically defined haptens which have one tenth the binding affinity of an Fab (Huston *et al.*, 1988b) and several examples of scFv's that recognise a complex protein structure with affinities only 2-4 fold less than that of their Fab counterparts. A caveat with some of these measurements is that they may have been influenced by the presence of incorrectly folded material (Pantoliano *et al.*, 1991).

Recombinant chimeric antibodies are based on the transplantation of murine CDRs onto human variable-chain frameworks and are intended as tools for diagnostics *in situ* and for therapy, since murine antibodies are recognised as foreign causing an immune response on secondary challenge. While the widely used method of achieving humanisation is the reshaping technology of Winter and Milstein (1991), there are additional alterations of individual amino acid residues within the framework that are usually necessary to restore full activity for antigen.

In a novel approach, Fv surface analysis of 12 Fab X-ray crystallographic structures has been used to exhaustively characterise the surface of different V-regions (Pedersen, 1993). This resulted in the development of a novel method termed Resurfacing. Three major observations were made; first, that for the residue frequencies at certain positions of the sequence were particularly conserved. Second that the amino acid positions identified none of the entire combinations of surface

residues in the human sequences were found in the murine sequences and *vice versa.* Third, only at two of the positions were different distributions of amino acids found (Roguska *et al.,* 1997). This method is now becoming widely used since in all instances so far tested full activity is retained after the resurfacing procedure (Roguska *et al.,* 1997).

Plant based expression systems present a cost-effective means for the productions of recombinant antibodies and proteins in general. It is possible to produce whole Fab and IgG molecules in a functional form in various organs of the plant (seeds or tubers). Plants have been shown to be capable of expressing functional engineered forms of antibody such as "single domain antibodies" (VH fragment alone) (Benvenuto *et al.*, 1991) or scFvs (Owen *et al.*, 1992 and Tavladoraki *et al.* 1993). These genes may have derived from hybridomas or from phage display libraries. This technology has implications not only for plant pathology but could be of general utility when other plant functions or activities are permanently modulated.

1.3.4 Gloop2: An anti-lysozyme antibody and a model for re-design

Hen egg lysozyme (HEL) is an ideal model antigen as its three dimensional structure has been determined to high resolution (Blake *et al.*, 1965). In addition, the structures of lysozymes from other species are known, for example human (Artimiuk and Blake, 1981), turkey (Bott and Sarma, 1976), hen (Smith *et al.,* 1993) and goose (Grutter *et al.*, 1983). A panel of primary sequences from closely related and more distantly related lysozymes is also available (Grutter *et al.*, 1983).

The antigenic response to HEL has been the subject of considerable study. Atassi (1978) proposed that the entire structure of lysozyme consisted of only three antigenic sites. However, subsequent studies have shown that most, if not all, of the surface of HEL is potentially antigenic (Smith-Gill *et al.*, 1984; Harper *et al.*, 1987). Structures of three anti-lysozyme antibodies have been solved by X-ray crystallography, D1.3 (Amit *et al.*, 1986; Bentley at al., 1989; Bentley *et al.*, 1990), HyHEL-5 (Sherrif *et al.*, 1987), HyHEL-10 (Padlan *et al.*, 1989). In addition, the structure of an anti-peptide antibody that cross reacts with native HEL, Gloop2, has been determined (Jeffrey, 1989). Gloop2 is one of a group of five mAbs (Gloop1 to 5) specific for the loop region of HEL, which comprises residues 57-84 (**loop peptide**). The antibodies were raised against a peptide-bovine serum albumin conjugate and cross react with native HEL (Darsley and Rees 1985a and b). The isotype of Gloop2 is γ2bk (Darsley and Rees 1985a). It binds the loop antigen with a dissociation constant (K_D) of $2x10^{-8}$M and HEL with a K_d of 10^{-7}M (determined by solid phase RIA; Roberts *et al.*, 1987). Serological mapping of Gloop2 using a panel of avian lysozymes indicated that the epitope on HEL comprises of the majority of the loop region and includes at least

eight residues (Darsley and Rees, 1985a). The epitope has been further refined by NMR studies of the Gloop1-loop complex (Cheetham *et al.*, 1991) - the epitopes recognised by Gloop2 and Gloop1 are virtually indistinguishable, as shown by Darsley and Rees (1985a). Studies on main chain mobility by NMR indicated that residues in the centre of the loop (e.g. Ser 72, Arg73) are not involved in direct contact with antibody as was originally assumed from the serological results. The epitope is now recognised to be discontinuous, consisting of residues 59 to 70 and 74 to 79 (Cheetham *et al.*, 1991).

The structure of Gloop2 Fv has been modelled by two different methods. The early molecular modelling studies (de la Paz *et al.*, 1986) used a maximum overlap method to build the combining site onto NEW framework, selecting the CDRs from available antibody structure according to length and sequence homology. The resulting model was then manually docked onto the serologically defined epitope of HEL and subjected to energy minimisation to relax any bad contacts between antibody and antigen (de la Paz *et al.*, 1986; Roberts *et al.*, 1987). This model was used to guide protein engineering experiments designed to test loop/HEL cross reactivity. Since then, the structure of the native Gloop2 Fab has been determined by X-ray crystallography (Jeffrey, 1989). Comparison of the structure with the model showed good agreement on four CDRs (r.m.s. difference 0.64 Å to 0.8 Å for the backbone). The other two CDRs, H2 and H3 were modelled incorrectly (r.m.s. difference on backbone 1.77Å to 3.61 Å) which prevented identification of a prominent groove in the centre of the combining site. Because of this difference in the modelled structure, the proposed docking with the HEL epitope was almost certainly incorrect. The Gloop2 crystal structure has been used for the development of novel modelling

methods which use a combination of both knowledge-based and *ab initio* algorithms. Two models of the combining site were generated, one in which each CDR was modelled individually in the presence of the other five (Martin *et al.*, 1989), and the other modelled *ab initio* (Martin, 1990). Both models had good agreement with the structure (r.m.s. difference 0.63Å to 1.36Å and 0.73 to 1.41 Å respectively).

The native Gloop2 structure was solved to 2.8Å resolution in two different crystal forms, enabling a comparison of two independently solved structures (Jeffrey, 1989). The combining site structures are similar with the only significant difference being the side chain positions of Phe52 and Tyr59 in H2, although this could be explained by the crystal packing interactions (Jeffrey, 1989). At the centre of the combining site there is a groove 12Å long, 9Å wide and 7Å deep which is defined by L3, H2 and H3.

The cDNAs for Gloop2 heavy and light chain have been cloned and a mutagenesis system was developed (Roberts and Rees, 1986). The mutant antibodies were expressed in *Xenopus laevis* oocytes after microinjection of synthetic RNA. The secreted antibody was tested directly for antigen binding affinity. Results from the early mutagenesis experiments show that removal of two charged side chains from the combining site (Glu28Ser (L1) and Lys56Gln(H2)) increased the affinity to HEL by 10 fold, making it equivalent to the affinity for peptide (Roberts *et al.*, 1987). Since then, extensive mutagenesis has surveyed the complete area of the combining site in order to define regions crucial for crossreactivity between peptide and native HEL (Roberts, S. pers. com.).

The *E. coli* expression system for Gloop2 Fv has been developed for the production of antibody fragments for structural analysis by NMR and crystallography (Field, 1988; Field *et al.*, 1989). The VH and VL polypeptides were expressed separately as insoluble cytoplasmic products. Partial purification and refolding *in vitro* to form the Fv heterodimer in large losses of material (90%) (Field *et al.*, 1989).

The Gloop2 system has been well defined by structural techniques (such as X-ray crystallography and NMR), molecular biology and protein chemistry. It therefore represents an ideal system to study antibody function and to attempt to create antibody molecules with novel properties.

1.4 Antibody Design

1.4.1 Modelling and Design of the Antibody Combining Site

The aim of antibody modelling is to be able to suggest a three-dimensional model corresponding to a given sequence.

Modelling of the antibody combining site was first attempted by Padlan and Davies (1976) at a time when very few antibody structures were known. It was recognised that the key lay in the structural homology that existed within the β-sheet framework regions of different antibody variable domains. Chothia's "canonical" families of residues for CDRs L1, L2, L3, H1 and H2 (Chothia and Lesk, 1987) were useful for those CDRs that obey the rules exactly while further development was necessary for those CDRs that lack the appropriate key residues and also neighbouring CDRs that might hinder correct domain assembly of an otherwise canonical conformation.

In the *de novo* design the goal is to suggest a sequence of aminoacids which is able to bind a pre-defined antigen, using three-dimensional information. There are three different approaches to the design problem:

1. From the known three-dimensional structure of antigen search the database to define **surface matching** with complementary shapes.
2. Starting form a known antibody-antigen complex structure a homologous antigen structure is docked at the same orientation and **minimal changes** are made in the antibody in order to retain the binding.

3. Using small peptide/hapten antigens (**peptide *ab initio*** method) in complex with a large range of homologous compounds leads to mapping of the individual interactions assuming the paratope shape and orientation. This method, because of the small antigen size limits the number of interactions, increases the size of database of structures and facilitates the synthesis and use of many different analogues. It also keeps open the possibility of using proteins at a later stage.

1.4.2 The Design Process

The design process starts with a known antigen and an antibody of known structure that is chosen randomly. It is assumed that the specificity of anti-hapten antibodies will not be dependent on particular combinations of CDR lengths using the antigen molecular weight as a descriptor of the antigen shape. This length independency hypothesis is substantiated by crystal structure data that show the formation of similar grooves in binding sites by both long and short CDRs (Rudikoff *et al.*, 1981, Jeffrey P., pers. com. Webster *et al.*, 1994). For larger antigens (MW> 10,000) it is not possible to deduce anything from these data as the molecular weight does not accurately describe the shape of the binding epitope.

From the randomly selected antibody combining site the side chains that do not influence the backbone structure (non-canonical residues) or are not buried by the CDR backbone and framework are truncated to alanine residues and the resulting structure is termed "the alanine cushion".

The next step is to dock the antigen in a reasonable initial orientation in the combining site in order to obtain maximum interaction area and maximum satisfaction of

electrostatic interactions using a Monte Carlo simulated annealing method (Goodsell and Olson, 1990). After the antigen has been docked all the residue positions which can potentially interact with the antigen are reconstructed. Evaluation of this reconstruction is obtained by selecting only those residue positions for which a favourable side chain rotamer exists and is capable of interacting with the antigen. Otherwise the original sidechain is retained. All sidechain conformations for each position are ranked based on energy data of the sidechain conformers and the antigen alone. By this procedure the sidechains that interact best with the antigen, and have the lowest electrostatic interactions are scored highest (Pedersen, 1993 Fig. 1.11).

The final residue rank for each sequence position is obtained from the estimation of the binding energy of antibody-antigen complexes and the accessibility of the solvent surface area lost by the antigen (see formula below). According to Novotny (1991) the total free energy of binding (ΔG_{TOT}) is a function of the hydrophobic effect (ΔG_F) upon antigen binding, the electrostatic interactions (ΔG_{EL}), the loss of sidechain conformational entropy upon antigen binding ($T\Delta S_{CF}$), the loss of overall rotational and translational entropy ($T\Delta S_{TR}$) and a correction term that accounts for dilute concentrations of proteins in biological systems ($T\Delta S_{CR}$).

$$\Delta G_{TOT} = \Delta G_F + \Delta G_{EL} - T\Delta S_{CF} - T\Delta S_{TR} - T\Delta S_{CR}$$

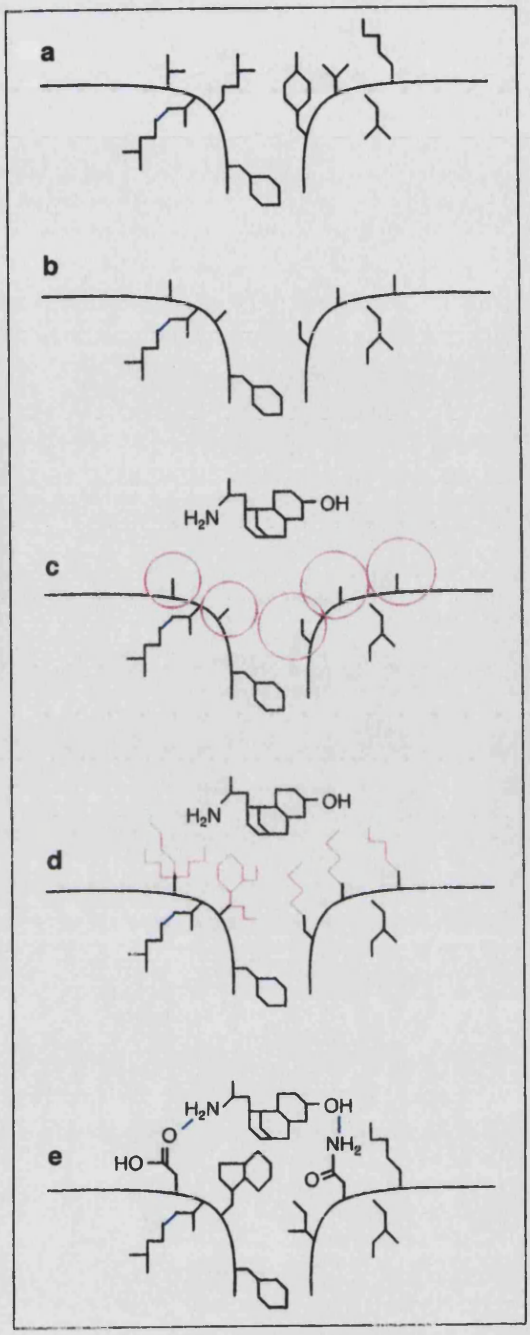


Fig. 1.11: The design process
Cartoon of the process described in the text by which an antibody specific for one antigen can be remodelled to fit a different antigen. In the example shown the new antigen is morphine, and the binding site is being created within the existing combining site of the anti-peptide antibody Gloop2. The process shown is as follows:

(a) the new antigen is selected for which an X-ray (and preferably and NMR) structure is known
(b) the side chains from all the CDR residues that are not required to define backbone structure are removed
(c) the 'alanine cushion' is generated
(d) and the new antigen is docked and the predicted contact chains are reconstructed
(e) the candidate side-chains are evaluated using knowledge based and conformational search methods. (Hydrogen bonds are shown in blue) (Rees *et al.*, 1994).

1.5 The aims and scope of this thesis

The scope of this thesis is to test experimentally the *ab-initio* design of antibodies (changing specificity) using the anti-peptide antibody Gloop2 (raised against a peptide, part of lysozyme, (Geffrey *et al.*, 1991) as a model system. The Gloop2 combining site has been re-designed using a genetic algorithm and computational methods to form a new, anti-opioid antibody GlaMor (Pedersen, 1993).

The design process presented in Chapter 3, predicted residues candidate for mutation that would change the Gloop2 specificity resulting in a novel anti-opioid combining site (GlaMor). Starting with Gloop2 and the set of mutations the experimental test had a dual aim.

First, to create the repertoire of mutations as phage displayed antibody fragments and select the "best" combination of these mutations on the basis of specificity against the opioid enkephalin. The Recombinant Phage Antibody Technology, described in Chapter 4, was used as a tool for this first test of the design.

Second, to express Gloop2 and GlaMor in a soluble form in *E. coli* and again test and compare their specificity for the lysozyme peptide and enkephalin respectively. In Chapter 5 different vectors are tested for their efficiency to express soluble form of scFv antibody fragments. The periplasmic expression in *E. coli* and purification of scFv antibody fragments together with optimisation studies is described in Chapter 6.

At a rather late stage, the Fv sequence selected on the basis of its binding to enkephalin as a phage construct when subjected to expression as a single chain Fv showed signs of significant instability. The analysis of this Fv and the attempts to obtain protein from it, are discussed in Chapter 6 and further extensive attempts to correct its sequence and expression problems are discussed in Chapter 7.

Finally Chapter 8 presents the qualitative assays for testing the binding properties of antibody fragments exploiting the Surface Plasmon Resonance technology.

2

Materials and Methods

This chapter describes the standard methods that were used as well as any modifications made, while more specialised protocols are discussed in the relevant chapters.

2.1 Materials

2.1.1 Bacterial strains

Escherichia coli strains used are listed in table 2.1. Strains like XL-1 Blue, DH5a and JM109 were used routinely for cloning and general DNA manipulation while TOPP2 and TG1 strains were used for protein expression work.

Strain Name	Genotype	Reference
XL-1 BLUE	[F'::Tn10 (Tc^R) *pro*A^+ *pro*B^+ *lac*I^q *lac*ZΔM15] *rec*A1 *end*A1 *gyr*A96 (NaI^R) *thi* *hsd*R17 ($r_k m_k^+$) *sup*E44 *rel*A1 *lac*	Bullock, W.O. *et al.*, (1987)
DH5α	*end*A1 *hsd*R17 ($r_k m_k^+$) *sup*E44 *thi*$^{-1}$ *rec*A1 *gyr*A (NaI^R) *rel*A1 Δ(lacZYA-argF)$_{U169}$ (ϕ80lacZΔM15)	Hanahan, D., (1983)
JM109	*thiI* *rps*L *end*A^+ *sbc*B15 *hsd*R4 *sup*E D(*lac-pro*AB)/F'[*tra*D36,*pro*AB^+ *lac*I^q *lac*ZΔM15]	Yanisch-Perron, C. *et al.*, (1985)
TOPP2	Rif^r [F' *pro*AB *pro*AB^+, Tn10 (Tet^r)]	Hatt, J., *et al* (1992)
TG1	*sue* *hsd* DS *thi*S[*lac-pro*AB] F' [*tra*D36 *pro*AB^+ *lac*I^q *lac*Z ΔM15	Gibson, T., (1984)

Table 2.1: The different *E.coli* strains used in for DNA manipulations and expression studies.

2.1.2 Bacterial cell culture media

All cell culture media were purchased from Difco.

LB (1 Litre) 10g Tryptone, 5g Yeast Extract, 5g NaCl (BDH)

2 x YT (1 Litre) 16g Tryptone, 10g Yeast Extract, 10g NaCl

XL (1 Litre) 16g Tryptone, 15g Yeast Extract, 10g NaCl

2.1.3 Plasmid Vectors

All plasmid vectors used are listed in table 2.2.

Plasmid	Obtained from	Reference
pKK322	IGEN Ltd.	Amann, E. & Brosisus, J., 1985
p*Trc* 99A	Pharmacia	Amann, E. *et al.* (1988)
pUC119/ His6mycXba	Prof. R. Hawkins Dpt. of Oncology University of Bristol	Griffiths, A. *et al.* (1994)
fd-tet-DOG1	Greg. Winter, MRC Centre for Protein Engineering (Cambridge)	Hoogenboom *et al.* (1991)

Table 2.2: Plasmid Vectors used in cloning and expression studies

2.2 DNA Manipulations

2.2.1 DNA preparation

DNA for cloning, sequencing and transformation was prepared using the QIAprep Spin Plasmid Kit by QIAGEN Ltd. This protocol gives yields of 5μg of plasmid DNA from 5ml overnight cultures of *E. coli*.

Protocol:

1. 5ml cultures of *E.coli* grown overnight at 37°Cin LB (Luria-Bertani) medium were pelleted by centrifugation at 5,000 RPM, at room-temperature for10 min.
2. the pelleted bacterial cells were resuspended in 250 μl of buffer P1 and transfered to a microfuge tube.
3. 250μl of buffer P2 was added and the tube was inverted gently 4-6 times.
4. 350μl of buffer N3 was added and the tube was inverted immediately but gently 4-6 times.
5. the lysed cells were centrifuged in a minifuge at 13,000RPM at room-temperature for 10 min.
6. the supernatants were applied by pipetting on a QIAprep spin column already placed in a 2ml collection tube.
7. the columns were centrifuged in a microfuge at 13,000 RPM at room-temperature for 60sec and the flow-throughs were discarded.
8. the QIAprep spin columns were washed by adding 0.75ml of Buffer PE and centrifuged as in step 7.

9. the flow-throughs were discarded and the columns were centrifuged for an additional 1 min to remove residual wash buffer.
10. the QIAprep column was placed in a clean 1.5 centrifuge tube and the DNA was eluted by adding 50µl of H_2O to each QIAprep column, let stand for 1min. and centrifuged as in step 7.

Medium scale preparation of plasmid DNA

Medium sized preparations up to 1mg of plasmid DNA were obtained using the Wizard Maxiprep DNA Purification System (Promega). The method required 500ml of overnight culture and the protocol was followed exactly as described by the manufacturer.

Large scale preparation of plasmid DNA; LiCl Method

This method gives yields between 25-100mg of plasmid DNA and was used when large quantities of DNA were needed for cloning manipulations. The plasmid DNA prepared with this method is ultimately treated with RNase to remove any RNA contamination offers the purity and high yields of the CsCl preparation without the ultracentrifugation step.

Protocol:

1. 150ml of 2xYT medium containing appropriate antibiotic(s) and 2% w/v glucose was inoculated with a bacterial plaque containing the desired plasmid. The culture was grown at 37°C, 250 RPM overnight in a shaker incubator to an A_{600} of >1.5.
2. The cells were harvested at 4,000RPM for 15 minutes and the pellet resuspended in 50ml STE (100mM NaCl, 10mM Tris-HCl pH 8.0, 1mM EDTA) and centrifuged again as before.

3. The pellet was resuspended in 9ml of solution I (50mM Glucose, 25mM Tris-HCl pH 8.0), 1ml of fresh lysozyme (10mg/ml) was added and the suspension gently mixed for 2 minutes.
4. 20ml of solution II (0.2N NaOH, 1% w/v SDS) was added and the suspension gently mixed and incubated at room temperature for 10 minutes.
5. 10ml of ice-cold solution III (4M KOH pH 5.5, 11.5% v/v acetic acid) was added to the flask and the mixture inverted several times and incubated on ice for 10 minutes.
6. the lysed cells were centrifuged at 4,000RPM for 15 minutes with no brake to avoid disturbing the pellet.
7. the supernatant was filtered through gauze and 0.6 volumes of isopropanol was added and the solution left to incubate at room temperature for 10 minutes. This solution was centrifuged at 5,000RPM for 15 minutes at room temperature.
8. the supernatant was discarded and the precipitated DNA allowed to air dry. The pellet was resuspended in 2ml TE buffer (10mM Tris-HCl pH 8.0, 1mM EDTA).
9. 2ml of 6M $LiCl_2$ was added and the solution was incubated on ice for 10 minutes followed by centrifugation at 6,000RPM at 4°C for 10 minutes. The supernatant was decanted to a sterile COREX tube and the DNA ethanol precipitated (2.2.5).
10. the solution was centrifuged at 13,000RPM at 4°C for 30 minutes and the DNA pellet resuspended in 1ml TE buffer. 10µl RNase (10mg/ml in 10mM Tris-HCl pH 7.5, 15mM NaCl) was added and incubated at 37°C for one hour.
11. the RNase was inactivated by the addition of 50µl 10% w/v SDS and incubation at 75°C for 10 minutes. The DNA was ethanol precipitated and resuspended in 0.5ml TE buffer (2.2.5).

2.2.2 Agarose gel electrophoresis

DNA can be analysed using agarose gel electrophoresis followed by staining with ethidium bromide and visualisation by ultraviolet (UV) irradiation of the gel. The following protocol describes a typical minigel analysis; however, a variety of gel sizes may be used, depending on the specific application. TBE running buffer was used for ordinary DNA visualisation while TAE running buffer was used for DNA gel purification.

Protocol:

1. The minigel apparatus was assembled (spacers, comb etc.).
2. the required amount of agarose was preweighed and added to the appropriate amount of TAE or TBE 1X buffer in a flask. For example, to prepare an 1% agarose gel, 1.0g of agarose was added to 100ml of buffer.
3. the mixture was heated in a microwave oven for the minimum time required to allow all the agarose to dissolve. The heating was interrupted at regular intervals and the container was swirled to mix the contents. The solution should not be allowed to boil over.
4. the solution was allowed to cool to 50-60°C and ethidium bromide was added at a concentration of 0.5μg/ml before the gel was poured. The gel was allowed to form completely (typically, 30min at room temperature). The comb was removed from the gel, and a sufficient volume of TBE or TAE 1X buffer added to just cover the surface of the gel.
5. Blue loading dye 6X was added to the DNA samples before they were loaded into the wells

6. the gel apparatus was connected to an electrical power supply and an appropriate voltage was applied. For minigels, typical gradients used were between 1-5volts/cm
7. after electrophoresis was completed, the gel was removed and placed on a U.V. light box and photographed according to specifications recommended for the camera and film type used.

2.2.3 DNA Amplification; The Polymerase Chain Reaction (PCR)

PCR was developed in the mid-1980s, thanks to the discovery of *DNA polymerases* that are stable at high temperatures. PCR enables specified DNA fragments to be amplified from minute amounts of starting material (theoretically just a single molecule). The only limitation is that the region of DNA to be amplified must be already known or at least must be predictable with 90% certainty in the boundary regions where the *oligonucleotides* (primers) will anneal and prime the amplification.

Basic procedure for PCR amplification.

All PCR's were performed in the Perkin-Elmer, Model: PTC-100 thermo-cycling machine. Vent DNA polymerase (NEB) was used that ensures high fidelity amplification and utilises its 3'-5' exonuclease activity for proof-reading of the extension product.

Protocol:

1. The PCR reaction was set up in 0.2 ml thin-walled eppendorfs by adding:

- 5μl of a 10x stock of dNTPs (dATP, dCTP, dGTP, dTTP at 3mM)

- 5µl of NEB 10x ThermoPol buffer (100mM KCl, 200mM Tris-HCl pH 8.8, 100mM $(NH_4)_2SO_4$, 0.1% Triton X100)
- 0.5µl of BSA (10mg/ml) was 0.2ml thin-walled tubes.
- 1µl of each oligonucleotide primer (50pmoles/µl)
- 10ng of template DNA
- 0.5µl of Vent DNA polymerase (2 U/µl)
- Sterile Milli-Q water to 50µl total Volume

2. one drop of mineral oil was added to the tube to cover the reaction and prevent evaporation.
3. all PCR's performed included a *negative control* (no DNA template added) to test for the absence of contamination in the reagents and a *positive control* on a known template to test the performance of the buffers, enzyme, and the temperature cycle parameters.

Oligonucleotide primers

All oligonucleotides used were custom made by Perkin-Elmer and had been HPLC purified and lyophilised.

Successful amplification requires two primers about 20-30 nucleotides long that anneal stably to sites bordering the target DNA region that will be amplified. There were a few factors taken into consideration when designing an oligonucleotide primer.

a. The *length* of the primer should be a compromise between specificity and affinity. Longer primers raise the stringency of the PCR, and ultimately allow amplification with high specificity while shorter primers (e.g. 7-mers) might yield a number of different products. On the other hand, at a given stringency of annealing, shorter

primers should give fewest products because they will not be long enough to anneal stably to sites of low homology.

b. the base composition of the primers has to be as heterogeneous as possible, with an even mixture of all four bases. Whenever possible, an imbalanced distribution of G/C and A/T-rich domains must be avoided.

c. primers that are used in combination should have similar Tm values.

d. internal and self-complementarities of primers that could lead to primer-dimer" formation and reduce the efficiency of the PCR were checked using computer software (Amplify, Bill Engels 1992. University of Wisconsin).

Temperature cycling

Each cycle of PCR consisted of three separate steps:

a. *denaturation* at 94°C for 2min to separate any double-stranded DNA

b. *annealing (hybridisation)* of primers to DNA template for 2min.

For complementary oligonucleotides of 11-20 bases the annealing temperature was calculated using the Wallace equation: $T_m = 4(G + C) + 2(A + T)$ (Thein and Wallace, 1986). PCR optimisation experiments should start with annealing temperatures approximately 5°C above the calulated Tm.

c. *extension* of primer by Vent polymerase at 72°C.

1 minute per kilobase of expected extension product was allowed; longer extension times may lead to degradation of the extended product due to the polymerases exonuclease activity.

d. the reaction was cycled back between the annealing, extension and denaturing steps 28 times. The final step involved cooling to 72°C and maintaining this temperature for 2 minutes to ensure all extension products were double-stranded.

2.2.4 Restriction Enzyme Digestions.

Restriction endonucleases recognise and digest specific palindromic nucleotide sequences leaving a 5' protruding sequence, a 3' *protruding sequence* or *blunt ends.* Because of the palindromic nature of the restriction enzyme sites, fragments generated have *complementary* or *'sticky'* ends which ultimately can be linked (ligated) to vector DNA. Plasmid DNA and PCR products may be digested with restriction enzymes. PCR products with restriction sites placed very close to the end of the DNA may require varying amounts of flanking DNA around the recognition site. Therefore, if an oligonucleotide primer is designed with a cut site that is too close to the end of the DNA, the site may cut poorly or not at all after amplification. All digests were performed in the appropriate NEB buffer described in the NEB product catalogue. Digestion of plasmid DNA required enzyme at 1 unit per 1 mg DNA whereas digestion of PCR extension products required 3 units per 1 mg DNA. Digestions were generally performed at 37°C overnight and inactivated by heating at 65-80°C for 15-30 minutes.

Double Digestions

The reaction mixture should be appropriate for the restriction enzyme requiring the lower salt concentration of the two enzymes to be used. Some NEB enzymes required the addition of BSA to a concentration of 100μg/ml. If BSA was a buffer requirement for either enzyme, it was added to the double digest reaction since it does not inhibit any restriction enzyme. If the reaction buffers of both enzymes are similar, the DNA may simply be incubated in the presence of both enzymes.

Protocol:

1. 100-200ng of DNA was added in a 49µl total reaction volume, made up with 5µl 10x reaction enzyme buffer and sterile water.
2. 1µl of the restriction enzyme (diluted in water if high unit title) was added to reaction and mixed gently with the pipette tip.
3. the tube was incubated in a water bath routinely at 37°C (unless stated otherwise in the product guide) overnight.
4. the enzyme was then heat inactivated at 65-80°C for 15-30min and a 5µl aliquot of the reaction was analysed by gel electrophoresis.

2.2.5 DNA purification

B-agarase digestion of LMP agarose

The B-agarase digestion of LMP agarose was used as a general method for purifying digested vectors and DNA fragments prior to cloning. The digested DNA was analysed on an 1% (w/v) LMP agarose gel in TAE. The gel was cast and run at 4°C for 1hour at 40mA.

Protocol:

1. The DNA containing LMP agarose was cut out with a scalpel under low frequency UV light.
2. the agarose was incubated at 60°C to melt and 10% total volume of 10x Agarase buffer was added.
3. the reaction was allowed to cool to 40°C before adding B-agarase (1u/µl):

   2µl enzyme per 200µl 1% (w/v) agarose gel.

   3µl enzyme per 200µl 1.5% 9w/v) agarose gel.
4. the reaction was incubated at 40°C for 2hrs, 10%v/v of 3M NaOAc. pH 5.5 was added and chilled on ice for 15min.
5. the undigested agarose was spun at 13,000RPM at room temperature for 15min and the supernatant was decanted.
6. two Volumes of iso-propanol (pre-chilled at -20°C) were added and DNA was precipitated after 30min at -20°C.
7. the DNA was pelleted at 13,000RPM at 4°C for 15min and the iso-propanol was carefully aspirated off leaving about 50µl.

8. 1ml of 70% (v/v) ethanol was added and the tube was inverted 2-3 times and spun again as in step 7.
9. the ethanol was carefully aspirated and remaining traces of alcohol were left to evaporate on the bench for 10min.
10. DNA was recovered in 10µl of TE pH 7.5 or sterile H_20.

Spin-column chromatography

Spin-column chromatography was used post-PCR for the removal of primers, and between sequential reactions in cloning for buffer exchange and desalting. MiroSpin S400 columns (Pharmacia Biotech Inc.) are pre-packed with Sepharyl S-400 HR resin and equilibrated in TE buffer (pH 7.6).

Protocol (Based on the general protocol described in the Instruction Handbook):

1. The resin in the column was resuspended by vortexing, the cup was loosened and the bottom closure was snapped off.
2. The column was placed in an 1.5ml Eppendorf tube and pre-spun in a minifuge at 3,000RPM at room-temperature for 2min.
3. the column was placed in a new tube and the DNA sample was applied to the top-centre of the resin.
4. the column was spun as in step 2 for 3min. The purified sample was collected at the bottom of the support tube.

Purification of DNA by phenol extraction and ethanol precipitation

In many cases, DNA prepared from a bacterial cell extract contains significant quantities of proteins and RNA. The standard way to deproteinise the cell extract is to add phenol or a 1:1 mixture of phenol: chloroform. *Phenol* denatures and dissolves proteins leaving nucleic acid in aqueous solution. Samples of DNA which have been

treated with phenol are routinely then extracted with 24:1 (v/v) *chloroform: isoamyl alcohol* to precipitate remaining protein and reduce the amount of dissolved phenol which is in the aqueous phase. DNA in the aqueous solution can be precipitated by adding ethanol or iso-propanol in the presence of sodium or ammonium acetate and cooling to -20°C.

Protocol:

1. Equal volume of phenol, prepared as described in Maniatis (Maniatis and Sambrook, 1989) was added to the DNA mixture and vortexed gently.
2. the aqueous phase, which contains the DNA, was separated from the organic phase by centrifugation in the microfuge, at 10,000RPM for 1 min at room temperature.
3. the aqueous phase was transferred carefully into a fresh microfuge tube and an equal amount of chloroform: isoamyl alcohol (24:1) was added and step 2 was repeated.
4. in order to precipitate the DNA, 10% of 3M NaOH pH 5.5 and then 2 volumes absolute ethanol (chilled at -20°C) was added to the aqueous phase and the tube was incubated at -20°C overnight or for 30min at -80°C.
5. the precipitated DNA was pelleted by centrifugation in a microfuge at 13,000RPM at 4°C for 30min. The ethanol was removed with care and the pellet was dried in a desiccator for 10min.An extra wash with 70% (v/v) ethanol was included to remove excess salt from the pellet. The dried DNA was resuspended in sterile TE pH 8.0 or water, and stored at 4°C for further manipulation.

Spectrophotometric determination of DNA

Small quantities of DNA can be assessed by electrophoresing samples in an agarose gel with known standards, but samples from large-scale plasmid preparations were usually determined spectrophotometrically.

Protocol:

1. 5µl of DNA sample was transferred to 995µl of H_2O in a 1ml quartz cuvette, the contents were mixed thoroughly and absorbance readings were taken at 260nm and 280nm. The ratio 260:280nm absorbance ratio should be more than 1.9. If the ratio is less than 1.9 then contamination with protein should be suspected.
2. the quantity of DNA was calculated using the guide that 1ml of a solution with an A_{260} of 1.0 is equivalent to 50µg of double-stranded or 35µg of single-stranded DNA.

2.2.6 Ligations

Appropriately digested and purified vector and insert DNA can be linked together in a final ligation step prior to transfer to *E. coli* competent cells. DNA ligase covalently links, fragments of DNA to each other provided that a 5'-phosphate is in close proximity to a 3'-hydroxyl group. When protruding *complementary termini* (*sticky ends*) are present on the vector and insert DNA molecules they provide an obvious docking mechanism to bring the 5'- and 3'- ends of molecules together. However, ligation can also occur when there is no obvious mechanism for prior adhesion of fragments. *Blunt- end* ligation is not as efficient as the sticky-end ligation but it occurs when a higher concentration of DNA insert and ligase are provided.

Vector: Insert Ratio

After the vector and insert DNA have been prepared for ligation, the concentration of each was estimated by agarose gel electrophoresis along with molecular weight markers of a known concentration. In most cases, a 2:1 ratio of insert: vector was used for sticky-end ligations and a 10:1 ratio for blunt-end ligations. Approximately, 20ng of insert DNA (MW 750bp) was added to 50ng of digested vector (MW 3.5Kb) for sticky -end ligations. The amount of insert DNA was increased to 60ng for blunt-end ligations to ensure that enough 'ends' were present in the reaction mixture.

Ligation (using the T4 DNA ligase)

Using the 2:1 (or 10:1) insert: vector ratio the ligation reactions were set up as follows:

Protocol:

1. 1.5µl of 10x ligation buffer (0.5M Tris·HCl pH 5.5, 100mM $MgCl_2$, 100mM dithiothreitol, 10mM ATP, 250mg/ml Bovine Serum Albumin) was mixed with the digested DNA and 1.0 µl of T4 DNA Ligase (diluted 10x before use to 4 U/µl). The reaction volume was made up to 15µl using sterile H_20.
2. the reactions were incubated overnight at 16°C
3. T4 ligase was heat inactivated at 65°C for 10min before setting up the transformations
4. all 15µl of the ligation reactions were used to transform competent bacterial *E.coli* cells.

Protocol (using the Fast link ligation kit, Epicentre Technologies Co., Wisconcin):

1. 1.5µl of 10x ligation buffer (100mM NaCl, 0.1mM EDTA, 0.1% v/v Triton X-100, 1mM dithiothreitol (DTT), 50mM Tris-HCl pH 7.5) was mixed with 1.5µl of 10mM ATP, the digested DNA and 1.0 µl of T4 DNA Ligase (2 U/µl). The reaction volume was made up to 15µl using sterile H_20.
2. the tubes were allowed to incubate at room temperature for at least 1 hour after which the reaction was heated at 70°C for 15 minutes to inactivate the DNA ligase (Michelsen, B.K., 1995).
3. 10% of the reaction mix was used directly to transform competent *E. coli* cells.

2.2.7 Preparation of competent cells

In the early 1970's it was observed that *E. coli* cells that had been soaked in an ice-cold solution were more efficient at DNA uptake than unsoaked cells. A solution of 50mM $CaCl_2$ is traditionally used. Although the exact reason is not well understood a possible explanation is that $CaCl_2$ causes the DNA to precipitate on the outside of the cells or improves binding of the DNA in the cell well. The actual movement of DNA (transformation) into competent cells is stimulated by briefly raising the temperature to 42°C. Competent cells were prepared using the $CaCl_2$ method.

Protocol:

1. 10mls LB medium containing the appropriate antibiotics and 1% (w/v) glucose was inoculated with cells toothpicked from frozen glycerol stocks and grown overnight at 37°C.

2. 500µl of the overnight culture was transferred in 50ml fresh LB medium containing the appropriate antibiotics and 1% (w/v) glucose and grown at 37°C with shaking until OD at 550nm was 0.4.
3. the cells were centrifuged at 5,000RPM at 4°C for 5min.
4. the pellet was resuspended in 25ml ice-cold $CaCl_2$ (50mM) and incubated on ice for 20min and repelleted as in step 3.
5. the new pellet was resuspended in 5ml ice cold $CaCl_2$ (50mM)/15% (v/v) glycerol and dispensed in 200µl aliquots into pre-labelled Eppendorf tubes.
6. the aliquots were snap frozen in acetone/dry ice and stored at -70°C.

2.2.8 Transformations

Heat shock transformation was used to introduce plasmid DNA into competent *E.coli* cells. This method yields more than 10^7 transformants per microgram of supercoiled plasmid DNA which is satisfactory for most cloning procedures.

Protocol:

1. 100ng of plasmid DNA was added in 50µl pre-thawed *E. coli* competent cells and was incubated on ice for 1hour. When a ligation reaction was used to transform bacterial cells the amount of DNA was determined by the volume of the ligation (i.e. normally for ligations less than 20µl all DNA was used in the transformation).
2. the transformation reaction was incubated at 42°C for exactly 2min
3. 450ml ice-cold LB medium containing 1% (w/v) glucose was added to the heat-shocked cells and incubated at 37°C for 1hour.

4. 200µl of each transformation was plated on LB medium plates containing 1% (w/v) glucose and the appropriate antibiotics and incubated overnight at 37°C.

(For the ligation reactions the cells were centrifuged at 5,000RPM at room temperature for 5min, the pellet was resuspended in 200µl LB medium (1% (w/v) glucose) and plated as in step 4).

2.2.9 Screening for recombinant plasmids

PCR Colony Screening

Transformed *E.coli* cells toothpicked directly from a plate, were screened by PCR to identify clones carrying the recombinant plasmids. Different sets of primers were used depending on the vector and the cloned fragment.

Protocol:

Each toothpicked colony was transferred in 20µl sterile H_2O and pre-incubated for 10min at 94°C. 1 µl from each sample was added to the reaction mixture containing:

dNTP's	2µl
10x Buffer	2µl
Primer1 (10pM)	1µl
Primer2 (10pM)	1µl
BSA	0.20µl
DNA	1µl
H_2O	12.55µl
Vent DNA pol.	0.25µl
Total Volume	20µl

The reactions were cycled 30 times, programmed depending on the primers used and the DNA template size (see table and 2.2.3). 5µl aliquots from each reaction were analysed on an 1% (w/v) agarose gel.

In-well Lysis Screening

In-well screening (Epicentre Technologies Co., Wisconsin), adapted from Sekar (Sekar,V., 1987), is useful for identifying recombinant-plasmids based on differential migration of supercoiled plasmids with and without inserts in agarose gels. The only limitation is that the inserts must be of a sufficient size to allow discrimination based upon migration. Initially protoplasts are generated from colonies that potentially contain recombinant plasmids. The *protoplasts* are subsequently lysed in the wells of an agarose gel and supercoiled DNA is detected by ethidium bromide staining following separation by gel electrophoresis.

Protocol:

1. 15µl of Protoplast Buffer was pipetted at room temperature into a convenient number of wells of a round-bottom microtiter dish or microfuge tube.
2. the transformed colonies, 2-3mm in diameter, were removed from the agar plate using a sterile toothpick.
3. the toothpick was vigorously swirled in the well of the microtiter dish for 10-15 seconds, to deposit a sufficient number of cells.
4. the microtiter dish was incubated at room temperature for 10min to allow protoplast formation.
5. while the cells were incubated 8µl of lysis buffer was added into the appropriate number of wells of a submerged agarose gel.

6. the protoplast buffer containing the cells was carefully added to the wells of the agarose gel filled with the lysis buffer. The mixture was allowed to incubate in the gel well for 15min.
7. in separate wells of the gel, supercoiled molecular weight markers and a sample of the supercoiled plasmid (without insert) were added. The DNA molecules were separated by electrophoresis, stained with ethidium bromide and visualised. Samples containing plasmid with insert migrated more slowly that the plasmid without-insert control.

2.2.10 DNA Sequencing

Manual DNA sequencing (Sanger-Coulson Method)

Manual DNA sequencing was performed using the Sequenase kit v.2.0 (US Biochemical, Ohio) based on the chain-termination method (Sanger *et al.*, 1977).

The chain termination method requires approximately 1μg of a purified single-stranded DNA template,. Sequencing double-stranded DNA requires an extra step where the DNA template is denatured before being added in the sequencing reaction. The chain-termination method (or dideoxy-nucleotide method) utilises a DNA polymerase (Sequenase) to synthesise a radiolabelled complementary copy of the template DNA in the 5`-3` direction. Incorporation of a dideoxynucleotide terminated chain elongation selectively at A, C, G or T because of the absence of a 3`-OH group. This generates the ladder of oligonucleotides to be resolved on a high resolution denaturing polyacrylamide gel. The reactions are set up to include also a short synthetic oligonucleotide primer (homologous to a portion of the template) from which the DNA polymerase initiates synthesis, normal deoxynucleotides that are used to elongate the new strand and an ^{35}S-dATP to label the growing strand. The reactions are subjected to electrophoresis on a high resolution denaturing polyacrylamide gel and autoradiographed to visualise the DNA sequence.

Protocol:

Plate Preparation:

1. Two glass plates, a front plate 48cm long and a back plate 50cm long were cleaned and polished with ethanol.

2. the back plate was placed in a fume cupboard and covered with a layer of 20ml ethanol mixed with 0.6ml 10% acetic acid and 0.2ml Bind Silane (A-174). The plate was incubated for 5min and then rinsed with distilled water and polished with ethanol.
3. the front plate was treated with 5ml Repel ™ (dimethyldichloranesilane) solution for 5min and rinsed as in step 2.
4. the back plate was placed on a large beaker and the spacers were placed on the sides. The front plate was placed on top and the two plates were sealed together with yellow tape putting extra tape around the bottom of the plates.

Preparing and casting the gel.:

1. The acrylamide solution was prepared by mixing together 16ml 40% Acrylamide/bis solution (AccuGel), 33ml diluent and 5ml buffer.
2. when fully assembled including comb and bulldog clips 300ul of ammonium persulphate (16%(w/v)) and 25µl N', N', N', N'- tetramethylethylene-diamine (TEMED) was added to the mixture and mixed well.
3. the gel mix was poured between the two plates, the comb was placed upside down about 0.5cm from the top and the edges near the comb were clipped together with the bulldog clip. The gel was allowed to set for 1-2 hours.
4. once the gel had set tissue soaked in TBE (890mM Tris-borate, 890mM Boric acid and 20mM EDTA) was placed over the top opening to prevent the gel from drying out.

The sequenase reaction:

1. Double stranded DNA was denatured with NaOH and then the alkali was removed using Pharmacia Spin Columns (see chapter 2.2.5).

2. the annealing mixture containing the DNA, 2µl reaction buffer, 1µl primer (50pM) in a total volume of 10µl was incubated at 65°C for 2min, allowed to cool down at room temperature over 20min and then chilled on ice.
3. 2.5µl of each termination tube was added to A, C, G and T prelabelled tubes and incubated at 37°C before use.
4. the labelling mix was diluted 1:5 to working concentration.
5. 1µl DTT (0.1M), 2µl Diluted Labelling Mix 0.5µl ^{35}S-dATP and 2µl diluted Sequenase polymerase were added to the ice cold annealed mixture from step 2 and the labelling reaction was incubated for 2min at room temperature.
6. 3.5µl of the labelling reaction was transferred to each termination tube (A, C, G, and T) and mixed well. The termination tubes were incubated at 37°C for 5min.
7. the reactions were stopped by adding 4µl stop solution to each termination tube.
8. samples were stored at -20°C until further use.
9. immediately before loading on to the sequencing gel the samples were heated at 75°C for 2min.

Loading and Running the Gel.

1. the comb, the clips and the bottom tape were removed and the plates were attached to the sequencing apparatus.
2. the upper and lower tanks were filled with TBE and the gel was pre-warmed using constant power 45W, 1700V for about 30min.
3. just before loading the samples the comb was placed no more than 2mm in to the gel surface.
4. 5µl of each reaction were loaded in the appropriate lanes marked A, C, G and T.

5. the samples were run for 3-4 hours until the lower marker was near the bottom of the gel.

Autoradiography

1. The gel was removed from the apparatus and the plates were separated gently using a razor blade.
2. the gel, usually stuck to the back plate, was soaked in 10% methanol, 10% acetic acid for 5min and then in water for 15min.
3. the gel was dried on the plate in a drying cabinet at 80°C for 30min.
4. the plate with the gel was placed in a cassette with the X-ray film on top and allowed to expose for 2-3 days in the dark room at room temperature.
5. the film was processed using the automatic autoradiography processor.

2.2.11 Automated DNA sequencing - Fluorescent DyeDeoxy terminator method

For Automated sequencing, the samples were sequenced on a 377 DNA Sequencer ABI Prism (Perkin-Elmer) using the fluorescent dye terminator method at the *Automated DNA Sequencing Service*, School of Biology and Biochemistry, Bath University. The principle of this method is briefly discussed in Appendix I.

Protocol:

1. Samples to be sequenced were placed in 12µl total volume in a 0.2ml thin-walled tube containing 300-500ng of double-stranded DNA or 30-90ng of PCR product DNA (purified from an LMP agarose gel, 2.2.5) and 3.2pM of HPLC purified primer
2. The primer concentration was determined using the following formula:

$$\frac{pmol}{\mu l} = \frac{100(A_{260})}{(1.54nA + 0.75nC + 1.17nG + 0.92nT)}$$

   [n = number of A, C, G or T]
3. All DNA sequences were checked against the sequence chromatogram to ensure that each base had been correctly identified and called by the sequencing software.

2.3 Protein Methods

2.3.1 SDS PAGE Electrophoresis

All analytical electrophoresis of proteins was carried out in sodium dodecyl sulphate (SDS) polyacrylamide gels. In this technique the proteins are denatured and coated with a detergent by heating in the presence of SDS and a reducing agent. The gels can then be fixed, stained with Coomassie brilliant blue or used to establish a western blot.

Preparing and Running the Gel.

Protocol:

1. The glass plates were assembled together according to the manufacturer's instructions. The Mini-Protean II System (BIORAD) based on the method described by Laemmli (Laemmli, 1970) was routinely used.
2. a 12% w/v resolving gel was prepared by mixing 4ml 30% w/v bis-acrylamide solution (ProtoGel) with 3.40ml H_20 and 1.5M Tris-HCl pH8.8. 100µl of 10% w/v SDS, 40µl of 25% w/v ammonium peroxodisulphate (AMPS) and 4µl of N, N, N', N', - tetramethylethylene-diamine (TEMED) were added and briefly mixed.
3. 3.5ml of the gel was poured into the vertical plates of the Mini-Protean II system 300µl of water was carefully added to ensure a smooth interface between the stacking and resolving gel and the gel allowed to set.
4. A 6% w/v stacking gel was prepared by mixing 1.7ml 30% w/v bis-acrylamide solution with 6.7ml H_20, 1.25ml 1M Tris pH 6.8, 100µl of 10% w/v SDS and 40µl

of 25% w/v AMPS. 10µl of TEMED was added and the mixture poured onto the resolving gel, a comb was inserted and the gel was allowed to set.

5. samples were mixed with loading solution (50mM Tris-HCl, pH6.8, 25% v/v glycerol, 0.1% w/v bromophenol blue, 6% w/v SDS, 4% v/v β-mercaptoethanol) at a ratio of at 5:1 (sample: loading solution) and incubated at 95C for 5 minutes. Electrophoresis was performed in tank buffer (250mM Tris-HCl pH 8.3, 192mM glycine, 0.1% w/v SDS) at 200V until the bromophenol dye had moved to the bottom of the gel.

2.3.2 Staining SDS-polyacrylamide gels with Coomassie Brilliant Blue.

Proteins separated by SDS-polyacrylamide gels can be simultaneously fixed with methanol:glacial acetic acid and stained with Coomassie Brilliant Blue R250. The gel is submerged in a concentrated methanol/acetic acid solution of the dye, and excess dye is then allowed to diffuse from the gel during a period of destaining.

Protocol:

1. The gel was submerged in Coomassie dye solution (0.1% w/v Coomassie Brilliant Blue R250, 50% v/v methanol, 10% v/v acetic acid) and left to stain on a rocking machine at room temperature for 2-3 hours.
2. the stain was then removed and saved for future use.
3. the gel was destained in a destain solution (10% v/v methanol, 10%v/v acetic acid) as in step 2 for 1-2 hours or until the background was clear.

4. stained gels may be stored in water in a sealed plastic bag or wrapped in cellulose film (BIO-RAD) and dried under vacuum at 60°C.

2.3.3 Western Blot Analysis

In general, *proteins immobilised* on membranes are detected with antibodies in a three-step process. First, *the primary antibody*, an IgG directed against the protein in question, is added to bind to potential antigenic sites. In the second step, *a secondary antibody-enzyme* conjugate, which recognises general features of all IgGs (anti-IgG), is added to identify bands at which the primary antibody is bound. The enzyme (AP or HRP) conjugated to the secondary antibody catalyses a *colorimetric reaction* in the third step, when the appropriate substrate is added. The colour provides a visual indication of proteins recognised by the primary antibody.

Protocol:

1. Six pieces of 3MM filter paper (Whatmann) and one piece of Hybond-C Nitrocellulose membrane (Amersham) were cut to the size of the SDS-PAGE Mini-gel and soaked in transfer buffer (39mM glycine, 48mM Tris-HCl pH 8.3, 0.037% (w/v) SDS, 20% (v/v) methanol) for 10 minutes.
2. the SDS PAGE Mini gel after the end of the electrophoresis was briefly soaked in transfer buffer.
3. the proteins analysed on the mini gel were transferred (blotted) on the Nitrocellulose membrane electrophoretically using the 2117 Multiphore II Electrophoresis unit (LKB Bromma).

4. a typical set-up for Western Blotting includes in order from the anode plate (+): three pieces of the filter paper, one nitro-cellulose paper, the mini-gel, the three remaining pieces of the blotting paper and the cathode plate (-). Any overlapping paper larger than the gel was cut to gel-size with a razor blade and the apparatus lid was secured.
5. electrophoresis was performed at 80mA, 10 volts for 30minutes for one blot or at 160mA for two.
6. after blotting the nitro-cellulose was washed briefly in phosphate buffered saline (PBS: 108mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.75mM KH_2PO_4) and the Mini-gel was Coomassie stained to check the efficiency of the transfer (see section 2.3.2).
7. the membrane was placed in a glass dish and incubated in 30ml PBS, 2% w/v non-fat dried milk (Marvel) for 1 hour at room temperature to block non-specific absorption sites. The membrane was washed twice with 30ml PBS for 3 minutes and immersed in 30ml PBS/Milk containing 0.1% (v/v) primary antibody and incubated at 4°C overnight. Alternatively, the blocking step was performed overnight at 4°C and the membrane was incubated with the primary antibody for 2hs at room temperature.
8. the membrane was twice washed in 30 ml PBS, 0.05% (v/v) Tween-20 for 3 minutes and washed twice in 30ml PBS for 3 minutes.
9. the membrane was incubated in 30ml PBS/Milk containing 0.1% (v/v) secondary antibody (conjugated to horseradish peroxidase HRP) for 45 minutes at 4°C.
10. the membrane was washed as in step8.
11. the antigen-antibody complex was located by immersing the membrane in 60ml HRP developing solution (0.3% (w/v) 4-chloro-1-napthol, 8mM Tris-HCl pH 7.5,

125mM NaCl, 0.015% (v/v) H_2O_2) and incubation at room-temperature until an intense purple colour developed.

(All the incubations and washing steps were carried out under low agitation on a rocking apparatus).

3

Redesigning the Antibody Combining Site

The design process outlined in section 1.4 was tested using enkephalin/morphine as the target antigen and Gloop2 as the antibody scaffold. The work presented in this chapter has been described in the *Ph.D* thesis of Dr. Jan Pedersen (1993), although the text has been modified and added to in places.

3.1 The design of an opioid antibody (GlaMor)

3.1.1 Antigen selection

The opioids were chosen as the primary target for the design project out of three candidate molecular classes sampled from Cambridge Crystallographic database (Kennard, 1991). The three candidate molecular classes namely, non-peptide antigens caffeine and theophyllin, opioid peptide antigens and the non-peptide opioid morphine and helical peptides containing the sterically constraint amino-acid α-amino-isobutiric-acid (Fig. 3.1). The rationale behind these selections is to

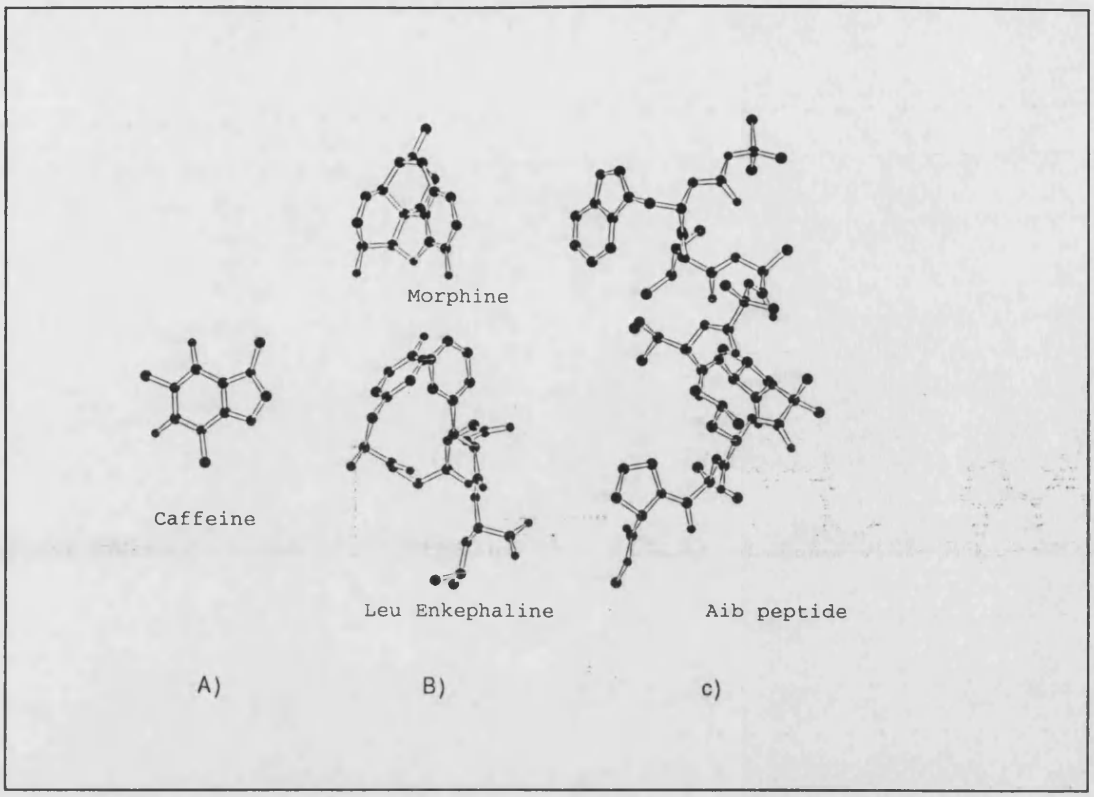


Fig. 3.1: Monte Carlo (Neumann and Ulam, 1945) plots of the three compounds chosen for the design an antibody.
The Monte Carlo programme is a general tool for molecular modelling targeting the prediction of side chain conformations in proteins. A. caffeine (Sutor, 1958a and b), B. morphine (Bye, 1976) and β-turn conformation of Leu-enkephalin (Aubry *et al.*, 1988), C. Aib helical peptide (Karle *et al.*, 1991).
The enkephalins, and related opiods have the most structural information available (From Pedersen, Ph.D thesis, 1993)

provide a spectrum of antigen types, a small non-peptide hapten, a small peptide hapten and a large peptide antigen.

The reason for choosing the enkephalins was that there exists the most structural information with structures determined by X-ray crystallography and refined to a resolution of <0.5Å (Aubry *et al.*, 1988). The N-terminal tyrosine residue of enkephalin -the one component of the enkephalin structure essential for biological activity-(see Fig.3.3) contains a phenolic ring separated by two carbon atoms from a nitrogen, which is also a feature of the morphine structure (Henderson and McFadzean, 1986).

Docking the antigen in the generic combining site as described in section 1.4.2 and finding the initial orientation requires data from known functional groups of the antigen. These date allow the determination of distance constraints for optimal liganding geometry.

The opiates have been extensively analysed in terms of their agonist/antagonist activities by QSAR methods (see Casy & Roberts, 1986 for a review); A pharmacophore model has been derived that has the elements shown in Fig.3.2.

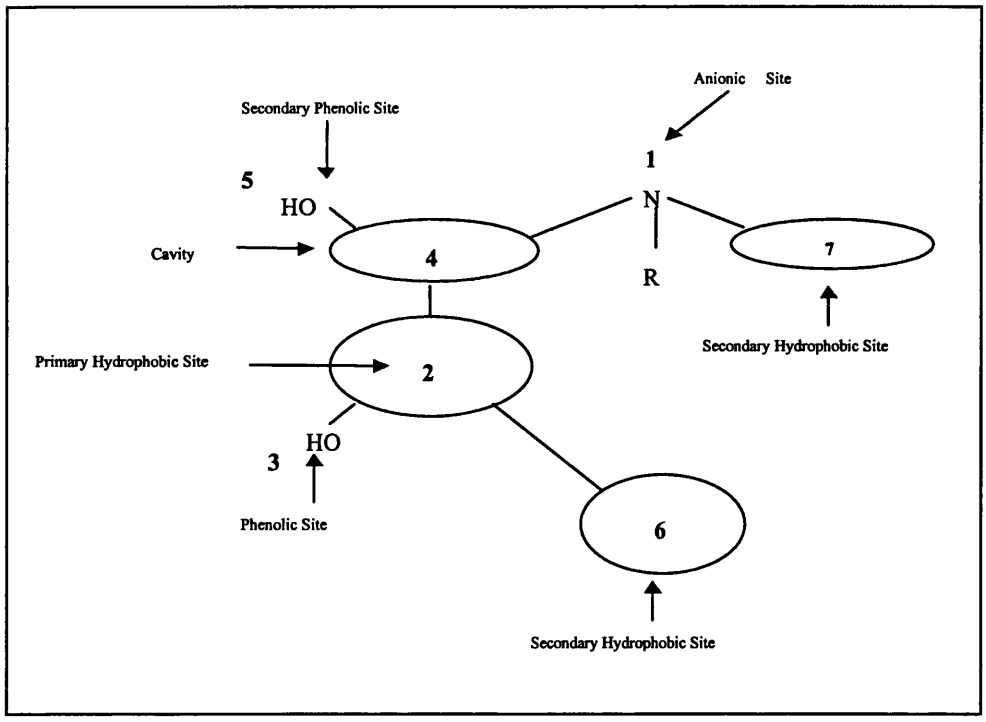


Fig. 3.2: Schematic Representation of the Opioid Receptor Binding Site.
Numbers indicate the putative positions of the liganding groups in the receptor. This mapping was the basis for the calculation of distance constraints (see text and Pedersen, 1993).

The putative receptor sites corresponding to the ligand features are indicated. Since the seven sights shown will have a defined distance relationship (a *sine qua non* of a pharmacophore model), the distance map derived from the pharmacophore was applied to each of the five opioid agonists, morphine, leu-enkephalin, naloxone, methadone and nalorphine. The structures of the agonists were obtained from the Cambridge Crystallographic Database (Kennard, 1991) and are shown overlapped in Fig3.3 (overlapping carried out using INSIGHT, MSI,USA). Average distances were then calculated and the final set of inter-site distances were used as constarints to search the combining site of the target antibody (see section 3.1.2) for potential complementary positions. Once found the selected set of seven contact positions would be mutated to a residue type most likely to provide the ideal complementarity to each of the opioid features. Since equally probable solutions for a given site might be allowed (given the inexactness of pairwise contact matching, either by energetics methods or docking), several sets of mutants at each position were generated.

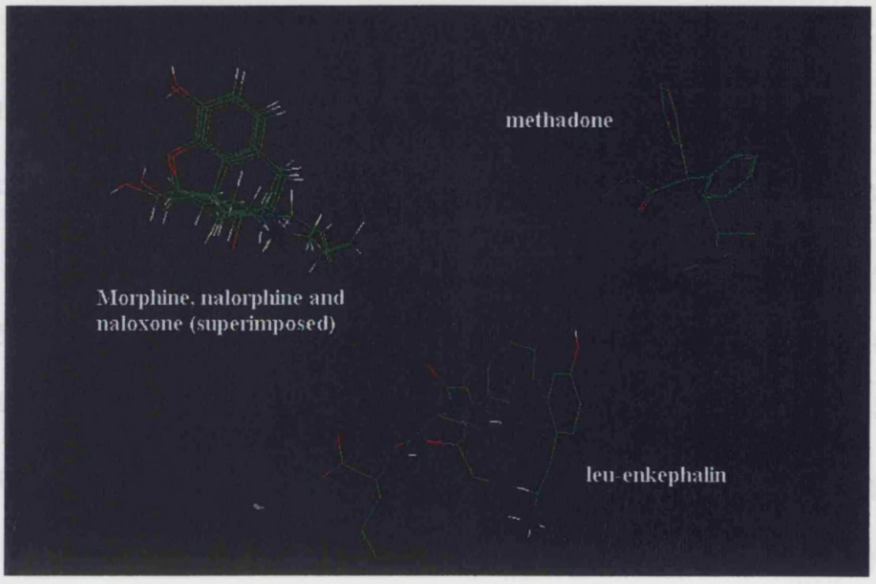


Fig: 3.3: Structures of morphine , nalorphine and naloxone overlapped, with leu-enkephalin and methadone in the likely equivalent conformation shown alongside

3.1.2 Antibody selection

The antibody selection process makes use of antigen topology information and the antibody classification as outlined in section 1.2.3. According to Webster *et al.* (1994) there are three groups namely cavity, groove and planar antibodies referring to the overall topography of the antibody combining site (Fig. 3.4). There are some indications (Rini *et al.*, 1992; Novotny, 1991; Herron *et al.*, 1989) that smaller antigens bind best when they are almost buried in the surface of the combining site, usually in a hydrophobic hole. Larger protein antigen prefer less curved surfaces and appear to bind over a larger surface area (Amit *et al.*, 1986 ; Sherrif, 1987), often with many charge-charge interactions.

MacCallum's recent classification (1996) clusters antibody combining sites into four topographic classes; concave and moderately concave (mostly hapten binders), ridged (mostly peptide binders) and planar (mostly protein binders).

The antibody platform chosen was the anti-lysozyme antibody Gloop2 (Darsley and Rees, 1985a and b). The specificity of this antibody for certain residues of the "loop" peptide of hen egg lysozyme was determined by serological mapping (Darsley & Rees, op.cit), by site-directed mutagenesis (Roberts et al 1987) and by NMR (Cheetham et al, 1991). The x-ray structure of Gloop 2, determined for the uncomplexed Fab only (Jeffreys, 1989), shows a truncated groove (see Fig.3.5a) formed at the centre of the combining site of dimensions ~12Å (long) by ~ 9Å(wide) by ~ 7Å (deep). It is lined predominantly by side chains although there

are a few contributions to the groove surface from main chain atoms. A list of the residues that line the groove is shown in Table 3.1.

CDR	Residue
L1	Y32
L2	A50, D55
L3	L89, Y91, L92, Y94, L96
H1	F32, G33, T35
H2	E50, F52, Y59
H3	E99, I100, R101

Table 3.1: List of residues of the Gloop2 groove

A

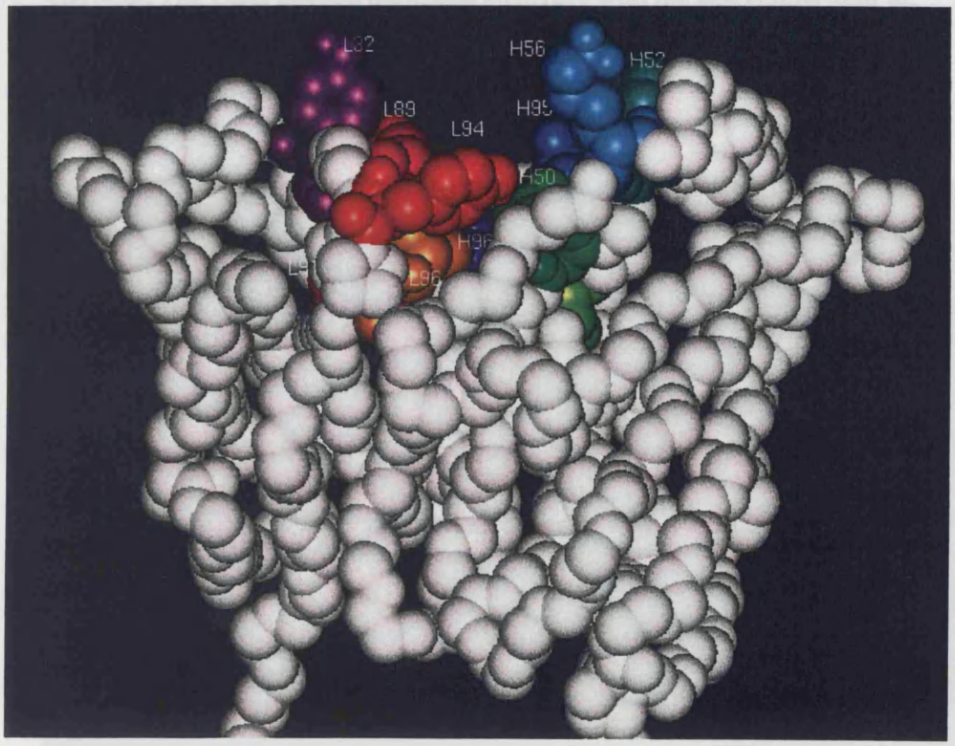


Fig. 3.5a: Space filling model of Gloop2 showing the groove type of combining site

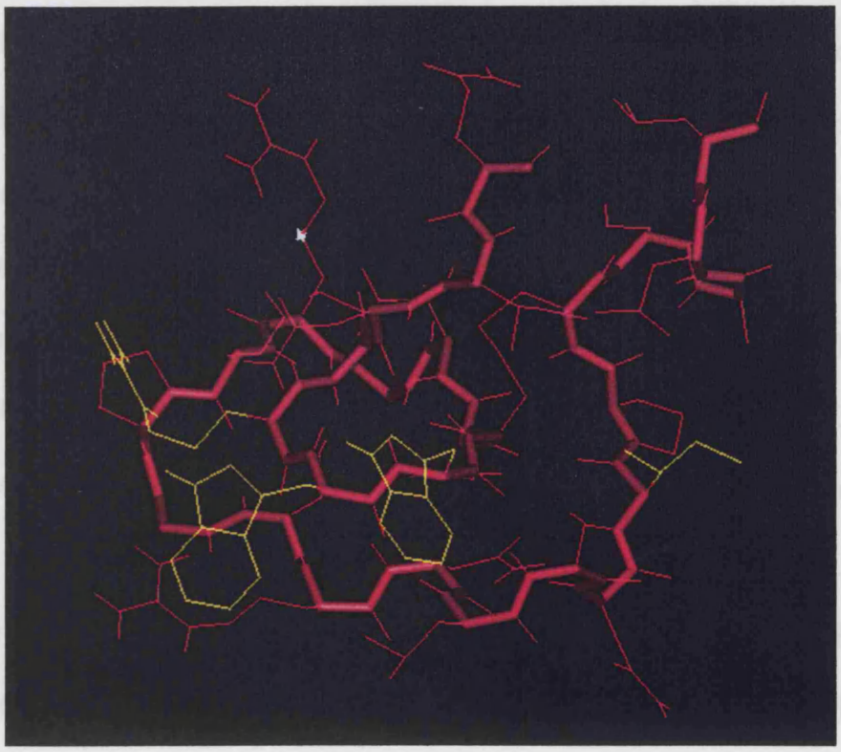


Fig. 3.5b: The Loop peptide backbone structure with the side chains built in.
The polypeptide sequence and the important residues (in bold) is shown below:
Gln-Ile-Asn-Ser-**Arg-Trp-Trp**-Cys-Asn-Asp-Gly-Arg-Thr-Pro-Gly-Ser-Arg-Asn-Leu-Ala-Asn-**Ile**-Pro-Cys-Ser-Ala-Leu-Leu

B

C

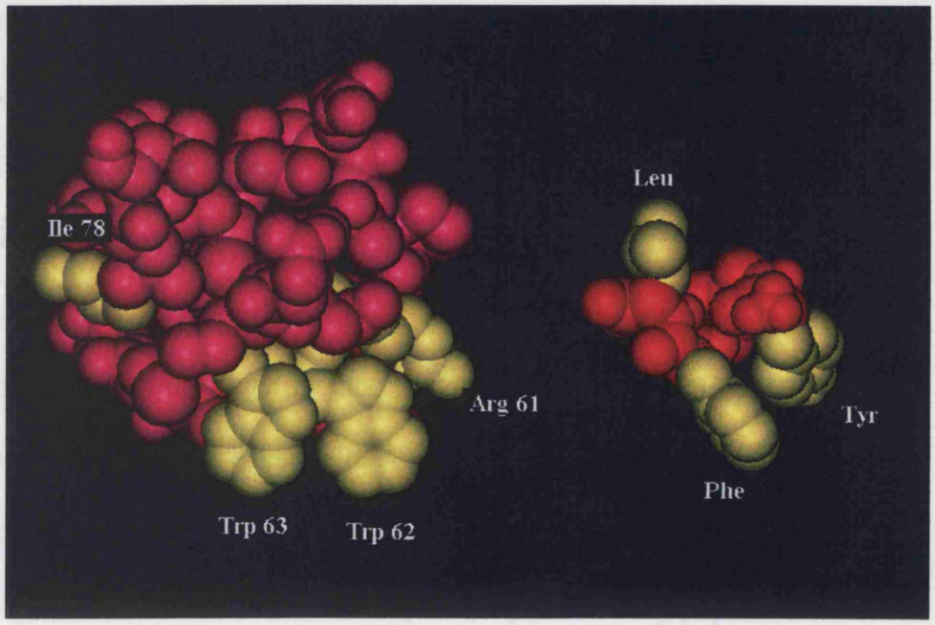


Fig. 3.5c: The initial (loop peptide-left) and the target (enkephalin-right) antigens in space filling models with the contact residues marked.

Of the six CDRs, L1 is almost completely excluded from the groove by L3 except for the sidechain of Y32, while L2 performs a secondary packing role. The groove of Gloop 2 is an extension of the hapten binding type pocket and is caused by the conformation of the unusually short H3 (4 residues). The Fab HyHEL-5 (Sheriff et al, 1987) has a small groove in this region also, except that it is terminated at both ends – by CDRs L3 and H3. Also, a fully formed groove is found in the NC41-neuraminidase complex (Tulip *et al*, 1989) which spans the combining site due to the packing of CDRs L3 and H3 against the VL and VH domains, respectively.

Of particular significance in the Gloop 2 groove is the presence of two partially buried Glu residues, E50 and E99. When mutated, binding to the loop peptide is lost, implicating charge-charge interactions in the binding of the parent antigen (Roberts & Rees, unpublished; Jeffreys, 1989). Confirmation of the likely role of the salt bridges in the Gloop 2-loop peptide interaction was obtained by NMR experiments in which line broadening (COSY experiments) was measured for each peptide residue during binding to the Fab (Cheetham et al, 1991). These studies clearly identified the roles of residues R61, W62,W63 and I78 in the interaction. The conformation of these residues in the crystal structure of lysozyme (Phillips et al, 1967) is shown in Fig3.5b where the loop peptide region only is shown; Residues implicated in the Gloop 2 binding are highlighted. For comparison, the structure of leu-enkephalin is shown alongside the loop peptide in Fig.3.5c. Of particular importance to note is that leu-enkephalin has no Arg residues, the roles of which are central to the specificity of the loop peptide for the Gloop 2 antibody.

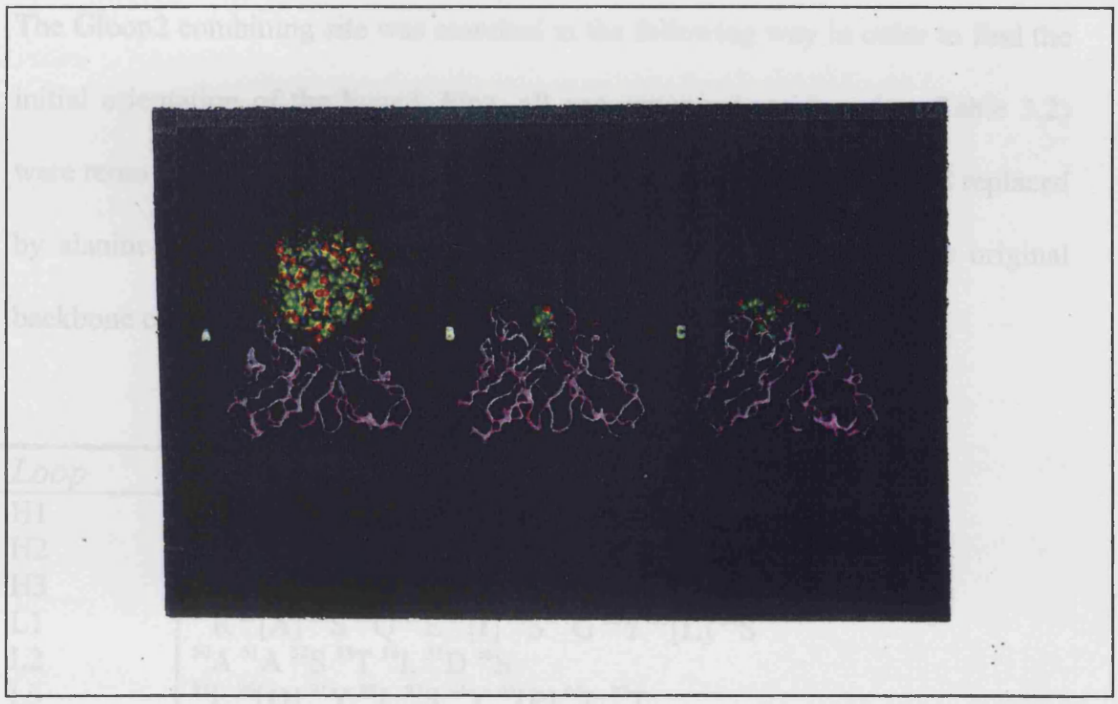


Fig. 3.4: The three antibody binding site types.
A. Planar, B. Cavity and C. Groove. In this picture the Fv framework is shown as the magenta ribbon, and the antigen is shown in a space filling model (Webster *et al.*, 1994)

Against this structural backdrop, the re-design of the Gloop 2 combining site was initiated. The groove of Gloop 2 was taken as the 'platform' and the following procedure adopted:

i) Map the positions of all canonical class 'key' residues in the CDRs (see Table 3.2) and retain the existing residue
ii) Convert the remaining residues to alanine
iii) Construct a distance map in which the distance constriants of the pharmacophore model (see Fig.3.2) are used to find the best fit to the Cß positions of the groove residues. This will now be described.

3.1.3 Search and Selection

The Gloop2 combining site was searched in the following way in order to find the initial orientation of the ligand. First, all non-canonical residues (see Table 3.2) were removed from Gloop2 CDRs' (retaining the canonical residues) and replaced by alanine, still accommodating the complementarity and retaining the original backbone conformation.

Loop	*Sequence*
H1	^{31}T ^{32}F ^{33}G ^{34}[**I**] ^{35}T
H2	[^{50}E ^{51}I ^{5}F ^{52a}P ^{53}G ^{54}N ^{55}S ^{56}K ^{57}T ^{58}Y]
H3	{^{95}E ^{96}I ^{101}R ^{102}Y}
L1	^{24}R ^{25}[**A**] ^{26}S ^{27}Q ^{28}E ^{29}[**I**] ^{30}S ^{31}G ^{32}Y ^{33}[**L**] ^{34}S
L2	^{50}A ^{51}A ^{52}S ^{53}T ^{54}L ^{55}D ^{56}S
L3	^{89}L ^{90}[**Q**] ^{91}Y ^{92}L ^{93}S ^{94}Y ^{95}[**P**] ^{96}L ^{97}T

Table 3.2: Canonical classification of Gloop2

Canonical residues are in bold and in [ ] brackets. For H2, all residues are in square brackets since at the time of the design, no canonical class was known. For H3, no canonical classes are known. Numbering is according to Kabat. (Kabat *et al.*, 1987)

A distance map of the Cβ positions was generated and the distance matrix was then searched against the distance constraints determined previously for the new antigen (actually, the pharmacophore). If a hit was found within the allowed variation of the distance constraint, the next constraint was searched for, and so on. If all the constraints were satisfied a hit was scored. The hits were ranked according to deviation from mean values obtained from the distances of the seven positions of the liganding groups in the opioid receptor binding site.

In the second search, all the non-canonical (or defined as being changeable) CDR residues (twenty six positions from Table 3.2) which could potentially participate

in specific ligand interactions were retained. These residue positions were then searched with the distance searching procedure described above. Identifying an initial orientation by this procedure reduces the number of residues that have to be changed and determines the maximum number of ligand requirements that can be satisfied by the original antibody sidechains.

In the starting orientation, where enkephalin was placed in the centre of the binding site in an orientation dictated by the constraint procedure above,, it was only possible to make contacts with thirteen of the original set of twenty six residues; The positions of those thirteen residues in the Gloop 2 structure are shown in Fig.3.6. After conversion of these thirteen positions to alanine, the enkephalin molecule was repeatedly 'thrown' into this site using a Monte Carlo algorithm (Neumann and Ulam, 1945) until a low energy set of docked structures was obtained (illustrated in Fig.3.7). Then, side chain replacement was carried out at each position and the lowest energy conformations of all the possible side chain types were extracted. The ten lowest energy residue combinations for all the residues in the construct are shown in Table 3.3, and formed the basis of the library to be constructed.

Sequence	Residue Position Number												
	L32	L89	L91	L94	L96	H32	H35	H50	H52	H56	H95	H96	H101
1	F	W	D	F	N	K	F	H	F	W	H	D	F
2	F	W	D	F	D	K	F	H	F	W	H	D	F
3	F	W	D	F	N	K	F	H	F	W	H	D	Y
4	F	W	D	F	D	K	F	H	F	W	H	D	Y
5	F	W	D	F	N	K	F	H	F	W	H	N	F
6	F	W	D	F	D	K	F	H	F	W	H	N	F
7	F	W	D	F	N	K	F	H	F	W	H	N	Y
8	F	W	D	F	D	K	F	H	F	W	H	N	Y
9	F	Y	D	F	N	K	F	H	F	W	H	D	F
10	F	Y	D	F	D	K	F	H	F	W	H	D	F
Original	**Y**	**L**	**Y**	**Y**	**L**	**F**	**T**	**E**	**F**	**K**	**E**	**I**	**R**

Table 3.3: The ten lowest energy conformations of the complete construct.
The original Gloop2 residues are shown in the bottom row

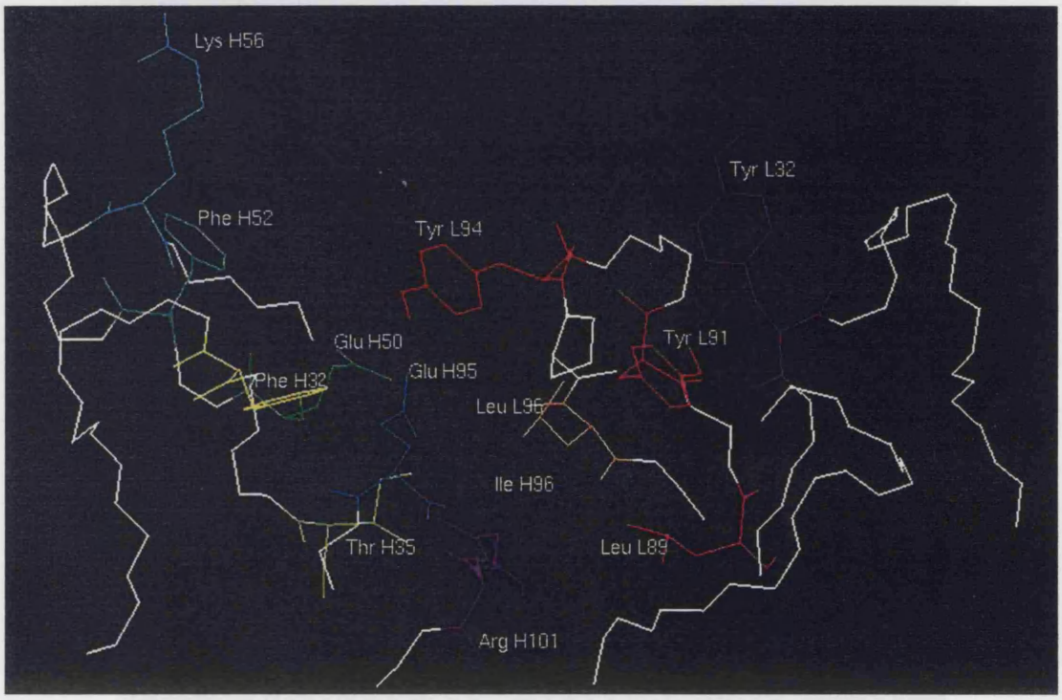


Fig. 3.6: The Gloop2 combining site with highlighted the proposed contact residues.

The heavy chain is on the left and the light chain is on the right.

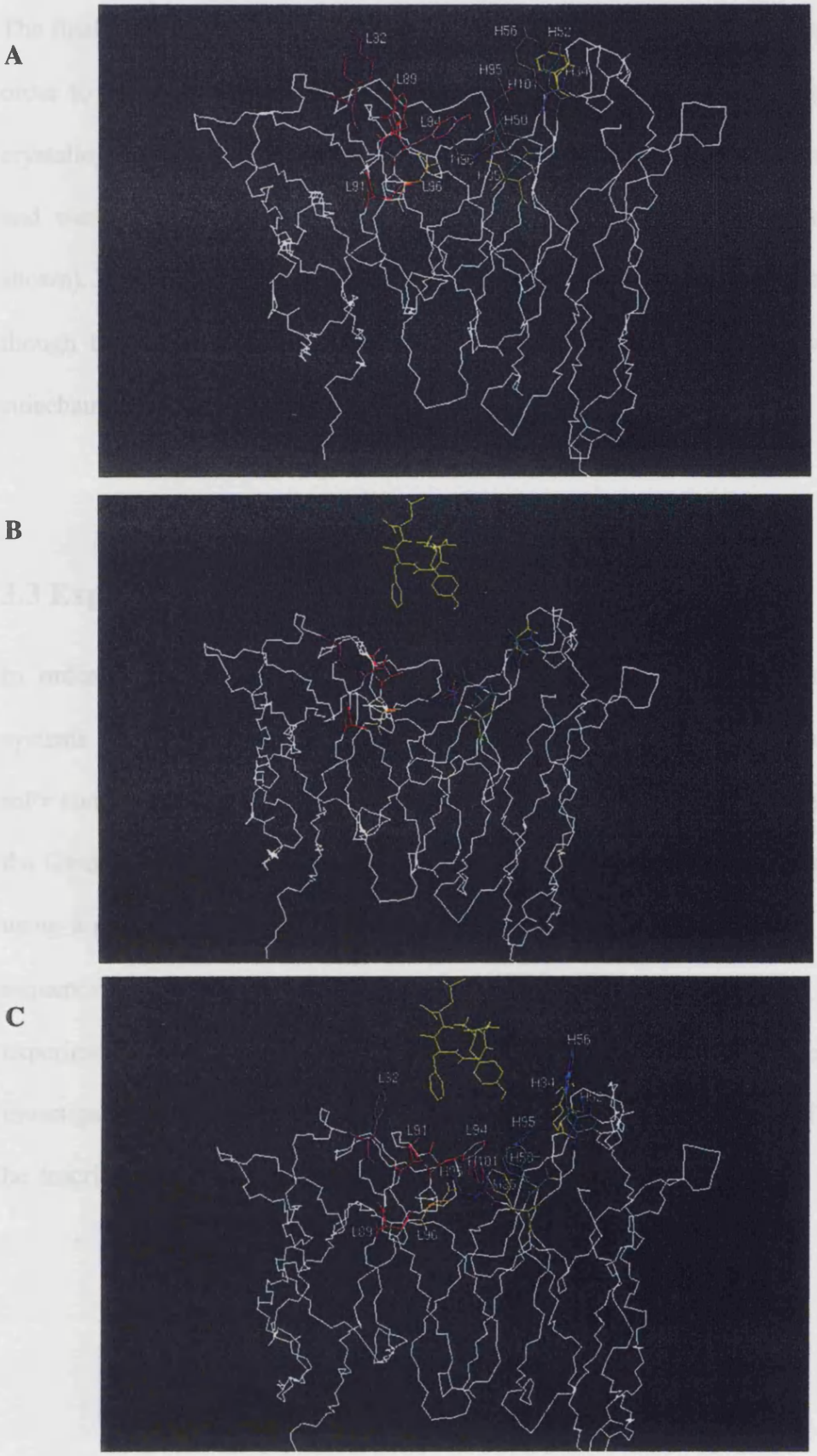


Fig. 3.7: Pictures showing the initial (Ala) and final binding site of the GlaMor antibody (Best construct)
Ala site, B. After side chain reconstruction (side view), C. After side chain reconstruction (front view).

The final ten best models were subjected to further energy minimisation cycles in order to validate the conformations. All the opioid ligands extracted from the crystallographic database were overlapped (Fig.3.3) in the GlaMor combining site and were able to assume the same orientation as the docked morphine (not shown). None of the non-peptide ligands clashed with the backbone of the model, though Leu-enkephalin exhibited some overlap of terminal sidechains with the sidechains of the reconstructed antibody combining site (Fig. 3.7).

3.3 Experimental test of the design

In order to determine how the design process performed, two experimental systems were devised and implemented in the laboratory. Gloop2 was built as a scFv construct. The ten GlaMor constructs were generated by PCR mutagenesis of the Gloop2 scFv, displayed on phage and screened for binding to leu-enkephalin using a phage display library. Binders from this library were sequenced and the sequences were compared with the ten predicted sequences. The second experimental strategy was to express soluble GlaMor scFv in *E. coli* and investigate its binding to leu-enkephalin using the Biacore technology. This will be described in the following chapters.

4

Phage display and selection of Gloop2 and GlaMor scFv

4.1 Introduction

Computational predictions on the effect of mutations on antibody affinity and specificity can be experimentally tested by mutagenizing residues in one or more of the six complementarity determining regions (CDRs) by site-directed mutagenesis. Using the M13 filamentous phage system, antibody gene fragments can be cloned in a phage vector expressed as a coat protein fusion and screened for novel antibody binding properties. Several libraries of phage displaying single-chain Fv fragments randomised at CDR positions have been constructed and selected for hapten affinity (Schier *et al.*, 1996).

The size of the library is directly influenced by the efficiency of transformation of the DNA into bacterial cells. Random libraries need to be of considerable size to contain a clone with the desired properties. If on the basis of structural information, randomising is restricted to a few residues, the number of possible mutations can be

adjusted so that the size of a library, covering all of them, is no longer a technical obstacle (Fig.4.1).

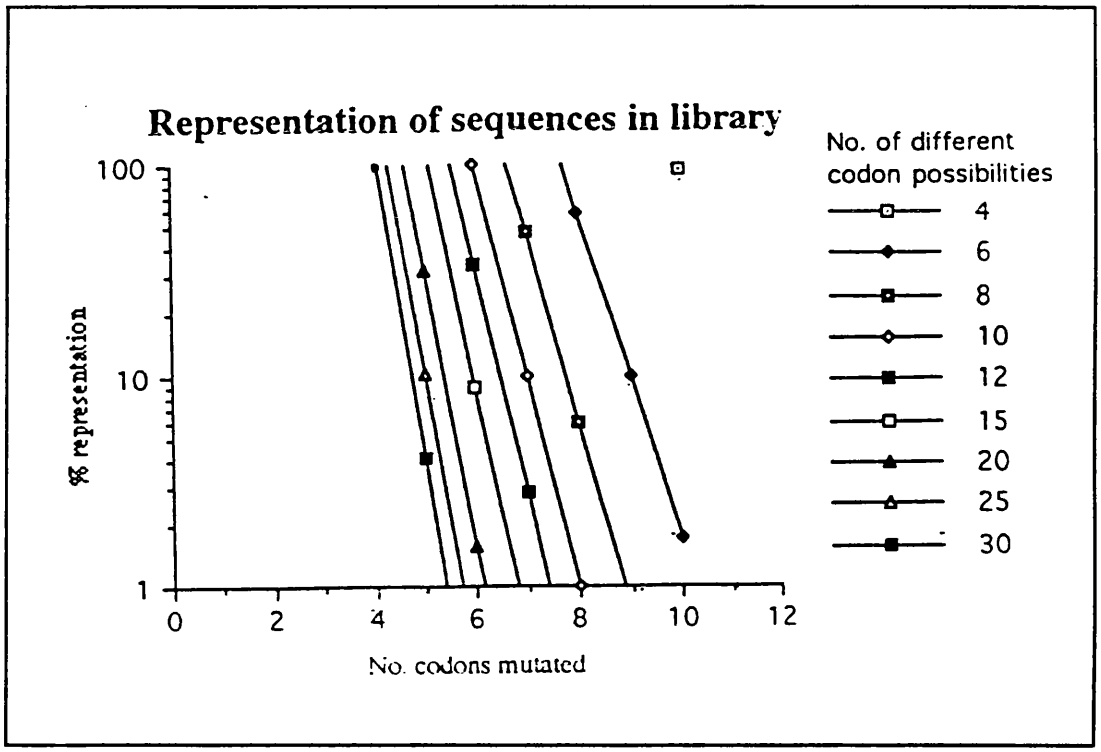


Fig. 4.1 Representation of sequences in the library
The amino acid positions on the protein sequence shown on X axis and the percentage of representation of different combinations of these mutations by a specific library size are shown on the Y axis.
Each of the labelled lines represents the different codon possibilities (amino acids) that are allowed. For example the -■- (first on the left) shows that if all 30 codon (all amino acids) are allowed then only 4 positions (30^4 combinations) can be mutated to get 100% representation (assuming the library size>30^4) and it requires $5x10^6$ transformants. If 5 positions are mutated then only 1% of the sequences will be represented by the above library size.

Antibody fragments are selected by subjecting the population of phage antibodies to antigen-affinity selection. Depending on availability of the antigen, and the type of binding desired, this can be accomplished by using columns of antigen immobilised on Sepharose (McCafferty *et al.*, 1990) or by 'panning' the library using plastic tubes or petri dishes which have been coated with antigen (Clackson *et al.*, 1991). After washing, binding phages can be specifically eluted, using for example an excess of antigen (Clackson *et al.*, 1991), or non-specifically eluted by changing the pH using,

for example, triethylamine (McCafferty *et al.*, 1990). Phages eluted from the affinity matrix are then used to infect *E. coli*. Since this particular phage genome carries a gene conferring drug resistance, infected cells can be selected on bacterial plates carrying an antibiotic. Phage antibodies are prepared from individual clones by growth in liquid culture and harvesting of the supernatant.

Phage antibody particles in the culture media can be used directly for subsequent analysis (e.g. by ELISA) and single- or double stranded DNA from selected clones can easily be prepared for sequencing or further mutagenesis. Soluble fragments can be produced without the phage particle by using vectors that contain an amber codon between the antibody and gene3. These vectors produce soluble or phage antibodies depending on the bacterial strain infected. Alternatively, soluble fragments can be produced by sub-cloning of the antibody genes into normal expression vectors (see Chapter 5).

4.2 PCR Mutagenesis

The approach of a restricted site-directed randomisation was used to design degenerate primers in order to introduce the predicted sets of mutations; Four positions in the sequence were randomised. The Gloop2 scfv gene was used as template. Gloop2 scFv (full sequence shown in Appendix I) was constructed by Dr. G. Elliott based on Gloop2 VH and VL genes (Roberts and Rees, 1986) and D1.3 scFv (generous gift from Dr. G. Winter, MRC, Cambridge) The VH and VL genes on the D1.3 scFv were replaced by the VH and VL genes from the Gloop2 Fv maintaining the linker sequence of the D1.3 construct intact. The sequences of the degenerate

primers synthesised with the mutations they carry are summarised in Table 4.1, while Figure 4.2a illustrates the basic mechanism of PCR site-directed mutagenesis.

Splice assembly PCR oligos for semi-random mutagenesis

For all reactions to generate fragment PCRs were performed with 30 cycles of 92°C (1min), 44°C (2min) and 72°C (2min).

Splice assembly PCR was performed with Vent DNA polymerase (strand displacement but no nick translation activity). Annealing temperature of the splice overlaps was 60°C for all fragments using the same programme.

Name	Sequence	No of bases	Mutations/ Sites
1FOR	ggt gca gct gca g	13	*Pst I* site
2BACK	cca ctc aag gcc ctg tcc agt tct ctg ctt cac cca **aaa** tat acc aaa ggt tgt g	55	T35F
3FOR	gga cag ggc ctt gag tgg att gga **cat** att ttt cct gga aat agt **tgg** act tac tac gct ga	62	E50H, K56W
4BACK	t ggt gcc ttg gcc cca gta g**(a/t)a** gt**(c/t)** atg tct tgc aca gaa ata g	44	R101(F/Y) I96(D/N), E95H
5FOR	act ggg gcc aag g	13	
6BACK	ttt ctg ctg aag cca gct taa **gaa** acc act aat ttc ttg ac	41	Y32F
7FOR	ctg gct tca gca gaa	15	
8BACK	ttt cag ctc gag ctt ggt ccc agc acc gaa cgt **gt(c/t)** cgg **aaa** act aag atc ttg **(c/g)(c/t)a** aca ata ata gtc tgc aaa	75	*Xho I* site L96(N/D), Y94F, Y91D and L89(W/Y/C),

Table 4.1: The primer sequences used for the site-directed mutagenesis.
The last column shows the mutations and the sites that are contained in each primer.
The library that will be constructed using the above primers will be semi-randomised at 4 positions and will contain 2x2x2x4=32 different sequence combinations.

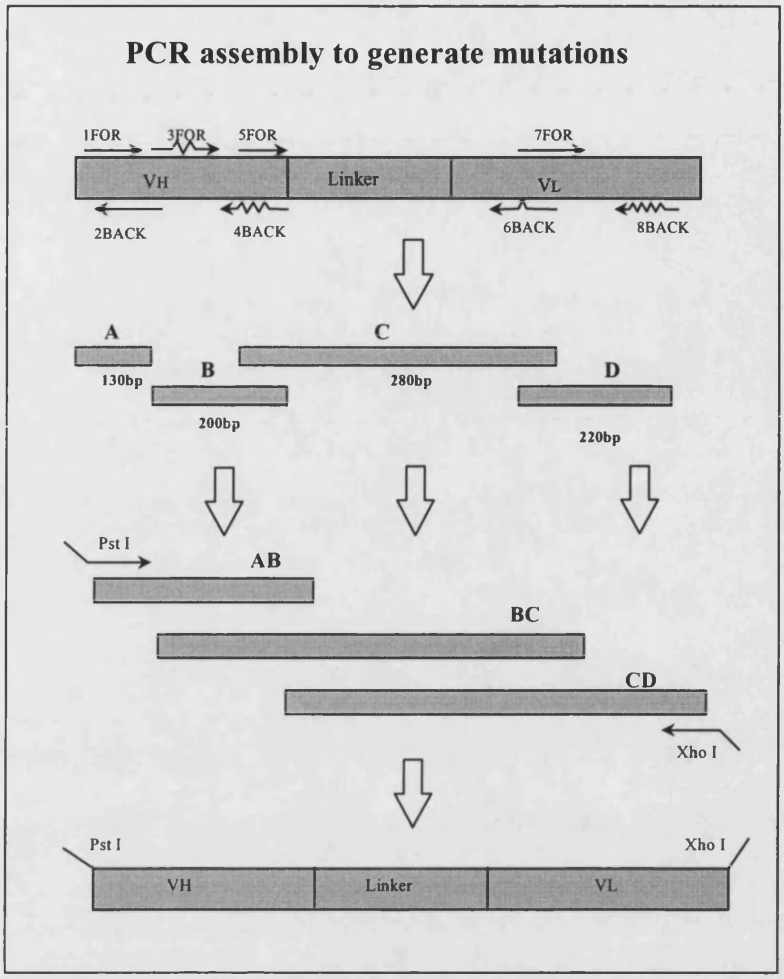


Fig. 4.2a: The basic mechanism of site-directed mutagenesis using PCR
Gloop2 scFv construct is shown at the top with the various degenerate oligonoucleotides (spiked arrows) used to introduce the mutations predicted by the Design process (Chapter 3). All possible combinations of mutations could be introduced resulting in the bottom scFv construct, which was cloned as a *Pst* I/*Xho* I cassette in the fd-tet-DOG1 phage vector.

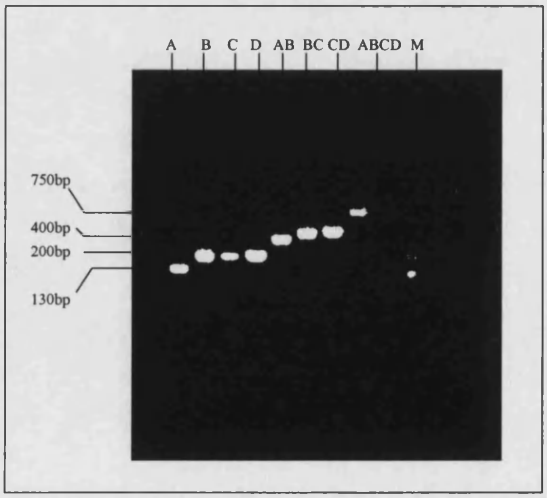


Fig. 4.2b: Agarose Gel of Mutagenised DNA fragments
DNA fragments after mutagenesis were analysed on an 1% (w/v) agarose gel. Sizes correspond to the ones described above. M; DNA markers

4.3 Library construction and multiple pannings

DNA Technology

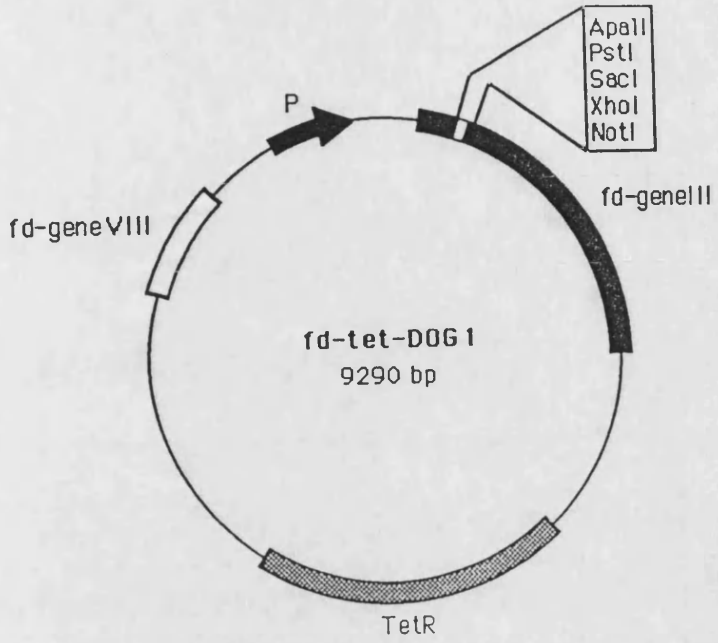
PCR products were cloned as *PstI* (in the VH)/*XhoI* fragments into the fd-tet-DOG1 (Hoogenboom *et al.*, 1991) for fusion with g3p (Fig. 4.3). Ligated DNA was electroporated (Dower *et al.*, 1988) into TG1 *E. coli* cells resulting in a mini library with 10^8 colonies/μg of DNA (Fig.4.5).

Titration of Phage

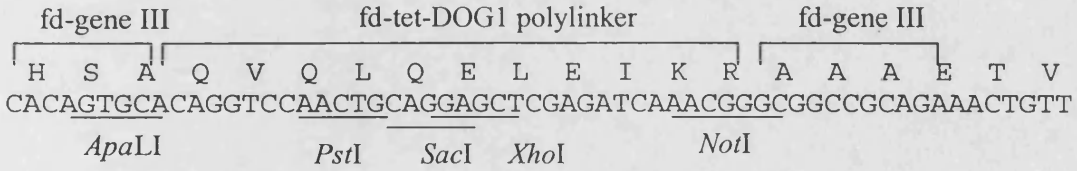
Protocol:

1. 10ml TG1 cell culture in 2xYT was set up overnight at 37°C
2. 100μl of the overnight culture was transferred in 10ml fresh 2xYT grown to log phase (OD_{600}: 0.5) and tranferred in 200μl aliquots to pre-warmed eppendorfs
3. serial dilutions of phage were prepared in water ranging normally between 10^{-2} to 10^{-14}
4. 50μl from each phage dilution was added to each aliquot of TG1 and incubated without shaking at 37°C
5. 200μl from each aliquot was transferred on 2xYT+15μg/ml Tetracyclin plates and incubated overnight at 37°C
6. colonies grown were counted and the original phage titre was estimated.

I



II



III

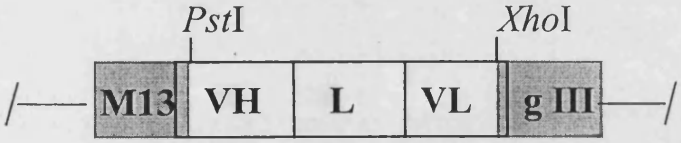


Fig. 4.3: Cloning the scFv genes in the fd-tet-Dog1 phage vector

I. Structure of the phage vector fd-tet-DOG1

II.. Sequence of fd-tet-DOG1 cloning sites

III scFv (Gloop or GlaMor) cloning cassette in the fd-tet-DOG1 phage vector

Panning library on streptavidin coated beads

Protocol:

1. Phage was diluted in sterile blocking buffer (1% casein/PBS/0.1% Tween 20) to a final concentration of 10^{9}TU/ml and preblocked by incubating for 1 hour at room temperature
2. Biotinylated enkephalin was added to a final concentration of 50ug/ml to 1ml of preblocked phage in a sterile eppendorf mixed on end-over-end shaker for 2hrs at 37°C. Meanwhile magnetic beads were preblocked:
    - Streptavidin coated magnesphere particles (Promega) were resuspended by vortexing. 0.5ml (0.5mg) of beads were transfered to an eppendorf tube and pelleted by spinning briefly (10sec, 2,000rpm). The tubes were placed into a DYNAL stand with the pellet towards the back. The supernatant was aspirated and the beads were centrifuged for 10sec at 2,000rpm to remove any residual preservative.
3. The beads were washed this way three times in 1ml of blocking buffer and finally were left to stand in 100µl of blocking buffer until the phage incubation was completed.
4. The beads were added to the phage/enkephalin solution and the incubation was continued for half an hour.
5. The beads were washed six times in 1ml PBS/Tween as before
6. The beads were resuspended in freshly prepared elution buffer (50mM triethylamine) and left rotating end-over-end for 10min. Meanwhile, for each pan 50µl of Tris HCl pH 7.4 was aliquotted into an eppendorf. The phage should not be left in the elution buffer for more than 10min.

7. The beads were pelleted, the phage supernatant was transferred to the Tris containing tubes and mixed immediately to neutralise.
8. Titre serial dilutions of 10µl aliquots were used to infect 200µl of exponential TG1, incubated for half an hour and then plated onto 2xYT+ 15µg/ml Tetracyclin plates. A PCR screen and/or sequencing of the resulting colonies was used to check the ratio of recombinant to parental vectors.
9. The eluted phages were reamplified by doing a large scale phage prep and then titrated as described above (Fig. 4.3).
10. Phages after three rounds of panning were used in a phage ELISA over a range of BSA-enkephalin concentrations, PCR screened and sequenced. The library panning and selection procedure is illustrated in Fig. 4.4.

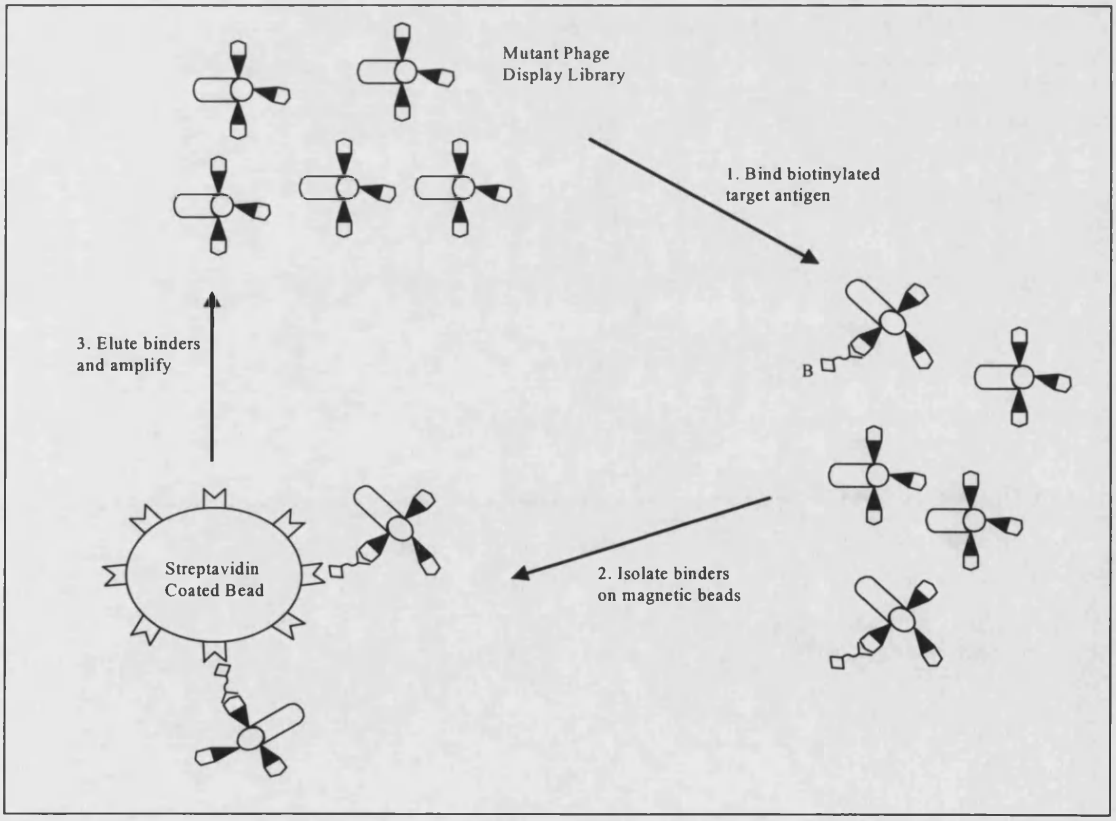


Fig. 4.4: Panning for anti-enkephalin mutants
Biotinylated enkephalin was added to phage from the mutant display library and any binders were captured on streptavidin coated beads. The binders were then washed, reamplified and panned again following the same cycle.

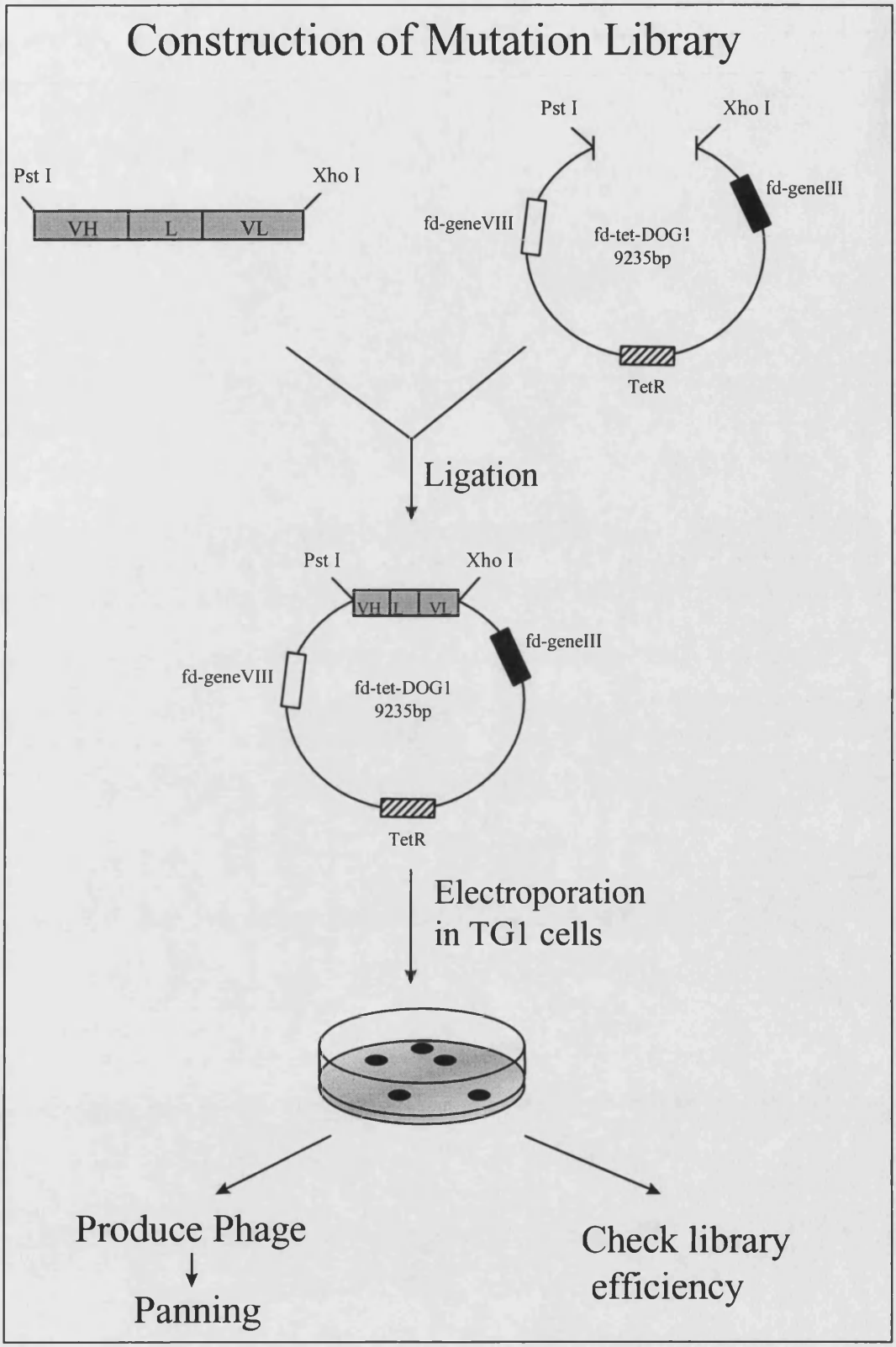


Fig.4.5: Construction of the mutation phage library.
The scFv DNA fragments (carrying different sets of mutations) were cloned into the fd-tet-DOG1 phage vector and electroporated into *E.coli* competent cells creating a phage library. Phage produced was panned against enkephalin to select for potential binders. The library efficiency was always checked to secure that all possible combinations of mutations were represented.

Preparation of Phage Supernatants, growth of cultures and reamplification of library Protocol:

1. Each colony was transferred to 10ml of 2xYT containing tetracycline and grown at 37°C overnight.
2. The cells were centrifuged at 13,000rpm for 5min a room temperature and the pellets were resuspended in 2xYT+15% (v/v) glycerol and stored at -70°C.
3. 5mls of the supernantant was filtered through an 0.45μm filter and stored at 4°C as phage stock

4.4 Phage ELISA

Protocol:

1. BSA-enkephalin conjugate was serially diluted over a range of 100ug/ml to 10ng/ml in fresh 50mM $NaHCO_3$ pH 9.6 using eppendorfs
2. 100μl per well of each dilution was pipetted onto a 96well Nunc maxisorb plate in duplicates
3. 100μg/ml of BSA was used a negative control and was put onto the end wells of the dilution series
4. The plate was incubated in a plastic box containing a little water overnight in the cold room. The plate was covered to prevent any drops of water falling in.
5. The samples were flicked out from the wells and washed by filling a sandwich box with wash buffer (PBS/0.1% Tween-20) and submerging the plate. The plate was tapped side to side to remove bubbles and then flicked to remove any wash buffer.

6. The plate was placed upside down on blue roll and tapped on the bench a few times to remove any remaining drops/bubbles. The plate was left to drain for a few minutes.
7. Each well was filled with 200µl of blocking buffer (3% BSA (w/v) /PBS/0.1%Tween-20 (v/v)) using a multichannel pipette.
8. The plate was covered and incubated in a humidified sandwich box at 37°C for 1 hour to block. For the last half-hour of the incubation, phage was preblocked.
9. Phage was diluted in blocking buffer to 10^9TU/ml of blocking buffer and incubated at room temperature for half-hour. Wild type Gloop2 phage and fd-tet-DOG1 were also tested as control samples.
10. The plate was rinsed by submerging in wash buffer and tapped dry. 100µl of preblocked phage was added in each well in duplicate for each row and for each phage type.
11. The plate was incubated at 37°C for two hours in a humidified container. For the last half-hour, HRP-linked anti-M13 conjugate was preblocked at room temperature. This was diluted 1/5000 with blocking buffer and 100µl was added to each well.
12. After the incubation, the plate was washed six times with wash buffer and drained by tapping onto tissue.
13. 100µl of preblocked HRP-linked anti-M13 conjugate was added per well and incubated for 1 hour at 37°C in a humidified box. Meanwhile 10mM sodium citrate solution pH 5 was made up
14. The plate was washed six times with PBS/0.1% (v/v) Tween-20 and once in citrate solution and was allowed to dry on blue roll

15. 100µl of developing solution (10mg ABTS in 45ml citrate buffer) was added to each well. Just before adding the developing solution to the plate, hydrogen peroxide was added (36µl hydrogen peroxide/42ml developing solution). After the colour has developed the plate was read at 414nm
16. The absorbance was plotted against the log of the conjugate concentration.

Results

10^{12} phage particles (titrated as described earlier) were prepared and panned against biotinylated enkephalin on streptavidin coated beads. Phage displaying the Gloop2 scFv were used as a control and were panned against the biotinylated loop peptide against streptavidin coated beads. Three rounds of panning were carried out each time checking the phage titre and the presence of Gloop2 mutants by a PCR screen.

Phage after the third pan were screened for specific binding against enkephalin by phage ELISA (as described earlier). The control assay was set up using Gloop2 scFv displaying phage. Both phage constructs (Gloop2 and mutant) were screened against both the peptides (loop peptide and enkephalin). Binding was determined with a chromogenic reaction between the anti-M13 HRP conjugate phage and its substrate and measurement of the absorbance at 414nm. The absorbance was plotted against the log of serial concentrations of the peptides (Fig. 4.6). The binding of the library and Gloop2 displaying phage against the loop peptide first, and enkephalin second was compared. Tighter binding against enkephalin was indicated by a clear shift of the binding curve of the library to the left (about 1 log) of the binding curve of the Gloop2 displaying phage. The binding curves against the loop peptide showed that Gloop2 phage bound tighter than the mutant library (about a 1.5 log shift). These results

indicated that the phage population in the library was expressing a construct that was responsible for the tighter binding. To identify the nucleotide sequence of this construct DNA from several clones was prepared and sequenced (as described later).

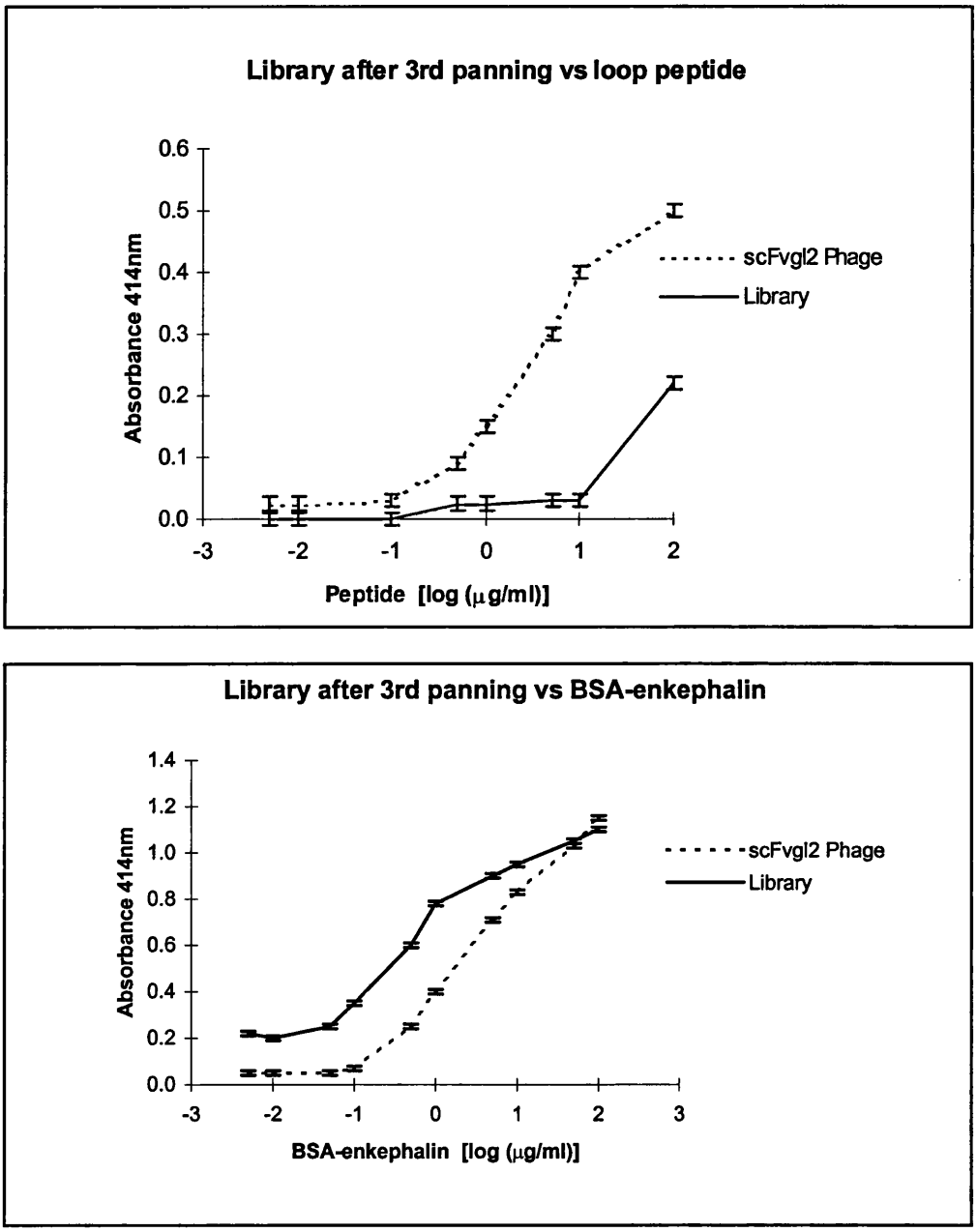


Fig.4.6: Specificity of the GlaMor phage library to enkephalin
After three rounds of panning and selection phage from the mutant library and Gloop2 scFv displaying phage were used in phage Elisa against BSA-enkephalin (bottom graph) and BSA-loop peptide (top graph). The peptide antigens, coating the wells, were serially diluted and were used to determine specific binding of the mutant display phage library to BSA enkephalin. The tighter binding of the library to BSA-enkephalin is shown with a shift of the curve to the left of the binding curve of the Gloop2 phage to BSA-enkephalin peptide.

4.5 Colony screening of panned library and sequences of positive clones

PCR aided colony screen

Eight colonies infected with phage after the third panning (titre:$2.58x10^{12}$) were PCR screened using oligos 5FOR and 8BACK to identify positive clones carrying the GlaMor scFv insert. Gloop2/fd-tet-DOG1 transformed TG1 was used as a positive control for the screen.

Each colony was toothpicked in a 30µl reaction set up with 0.6µl 5FOR (10pM), 0.6µl 8BACK (10pM), 6µl Buffer (5X) and 22.5µl H_2O and the reactions were heated for 10minutes at 94C before 0.3µl Taq DNA polymerase (Bioline) was added. PCR was performed with 25 cycles of 94°C (1min), 55°C (1min) and 72°C (2min) and then an extra step at 72°C for 10min. 5µl of dye was added to 10µl from each reaction and analysed on an 1% (w/v) agarose gel. Four colonies (50%) with scFv GlaMor inserts were identified and used to prepare ssDNA for sequencing (Fig.4.7a,b)

Preparation of ssDNA

Protocol:

1. 10ml overnight culture grown as above was used to innoculate 100ml of 2xYT containing 15µg/ml tetracyclin and was incubated at 37°C overnight
2. The cells were centrifuged at 3,000rpm for 10min at room temperature
3. The pellets were resuspended at 2xYT+15% (v/v) glycerol and stored at -70°C and 5ml of the supernatant was filtered using an 0.45µm filter and stored at 4°C

4. 1ml of 20% (w/v) PEG 6000 in NaCl was added to the rest 5ml of supernatant and incubated at 4°C for 1 hour or alternatively the cell pellet was resuspended in 1ml of 85% 2xYT (v/v) containing 15%(v/v) glycerol frozen on dry ice and stored at -70°C
5. The supernatants were centrifuged at 6,000rpm for 10min and the pellets were resuspended in 200µl of TE and purified twice using phenol chloroform and then ethanol precipitated. The final pellet was resuspended in 10µl of TE.
6. 3µl was analysed on an 1% (w/v) agarose gel and 7µl was added to the sequencing reactions

All four clones were sequenced manually using different sets of primers so as to obtain the full-length sequence. The sequences read were compared with the Gloop2 scFv sequence to identify any matches to the ten predicted mutation sets. All four clones gave the same sequence (Fig. 4.8 for the CDR sequences and Appendix I for the full-length scFv sequences) and this matched with the fourth ranked predicted construct but which contained the additional mutation I138R in CDRH1 (Fig. 3.2).

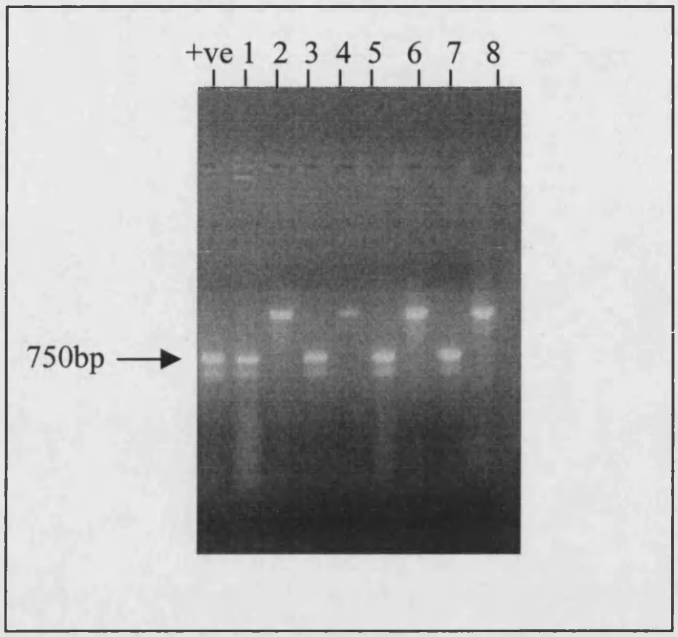


Fig. 4.7a: PCR screen of positive GlaMor/fd-tet-DOG1 clones
After the third panning eight colonies were PCR screened using oligos 5FOR and 8BACK. 10µl of each PCR reaction was loaded per well. Colonies 1, 3, 5,and 7 carry the right size construct that corresponds to the Gloop2 +ve control.

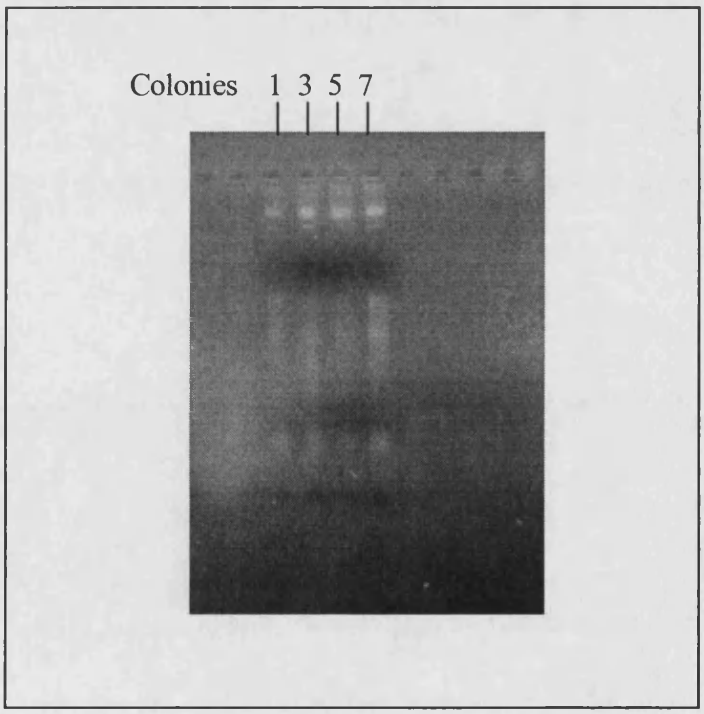


Fig. 4.7b: ssDNA preparation of mutant GlaMor/fd-tet-DOG1 positive clones
The four positive colonies (1, 3, 5, and 7) identified by the PCR screen were used to prepare ssDNA and used consequently for sequencing. 3µl of each DNA preparation was loaded per well.

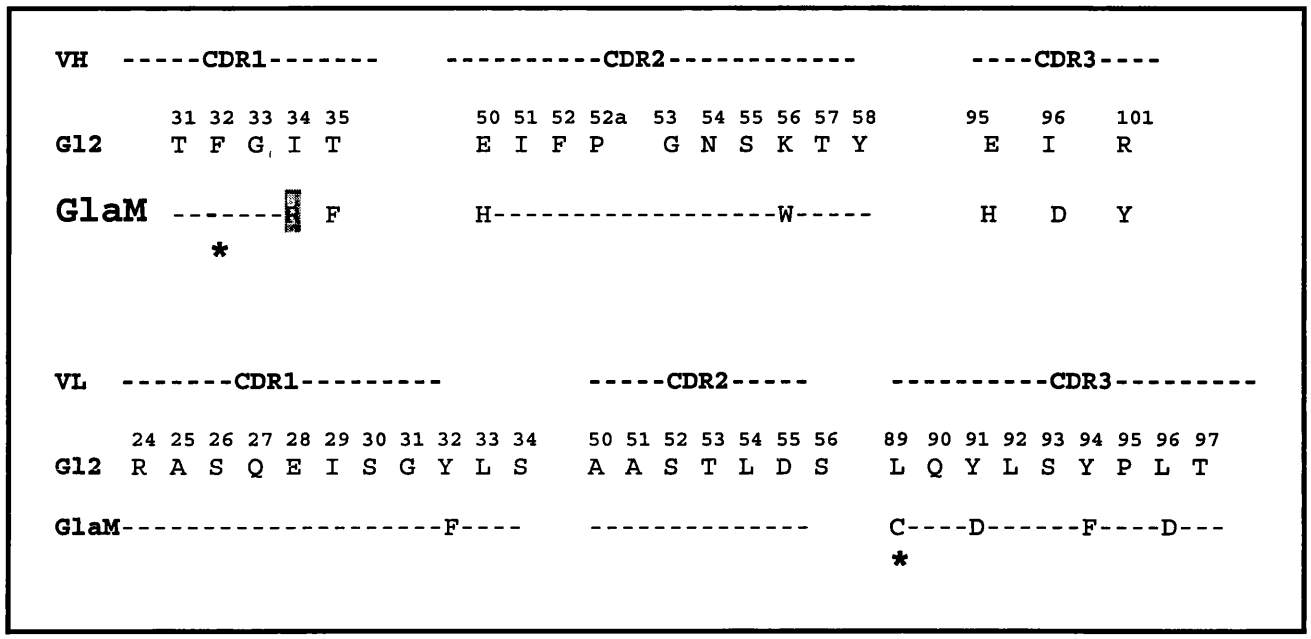


Fig. 4.8: Amino acid sequences of the Gloop2 and GlaMor complementarity determining regions in heavy and light chains.
Residues different between the fourth GlaMor design (see Table 3.2) and the phage-selected construct are marked with asterisks. The additional mutation is shadowed in grey. Residues are numbered according to Kabat *et al.* (1991). Details of the protein and DNA sequence of Gloop2 and GlaMor CDRs are described in Chapter 5 (Fig. 5.8a and b).

4.6 Discussion

The design process outlined in Chapter 3 identified twenty-six amino acid positions in the CDRs of Gloop2 combining site that were known not to be critical for CDR backbone conformation (CDRs L1-3 and H1). The absense of a canonical structure for CDR H2 (at the time of the design and construction) led to a conservative substitution for this CDR (2 positions only included), while three of the four residues in H3 were included in the design. From these 26 positions, only 13 survived the docking procedure (see Chapter 3). The all-alanine version of Gloop2 combining site and enkephalin was constructed and, after reconstruction of the side chains, twelve residues from these thirteen contact residues (7 in the heavy and 5 in the light chain) were taken as candidate positions for mutation. The resulting library with the different

combinations of these mutations was computationally screened, with a positive weighting applied to certain residues that occur with high frequency in antibody combining sites, and the ten lowest energy Glamor-enkephalin complexes were selected.

To assess the effectiveness of the design process, four of the twelve positions were semi-randomised (but also including the designed residue as part of the population). If one of the designed sequences was then selected during phage panning against a background of more diverse sequences, this would go some way towards validating the design process.

Mutagenesis of antibodies derived from phage display libraries has often given rise to drastic improvements in affinity (Marks *et al.*, 1992 and Schier *et al.*, 1996), as well as in neutralisation capacity (Burton, 1994). The advantages of generating a panel of mutagenised antibodies from a single starting clone was to fulfil the aim of assessing the *de novo* design of antibody specificity by defining a set of residues which is able to bind a pre-defined antigen (enkephalin).

There are a number of options available for introducing diversity into a clone prior to selection. Complete randomisation using degenerate oligonucleotides is restricted to a few residues by the library size governed by the efficiency of transformation in *E. coli*. While improved binders may be isolated, it is not always possible to pinpoint the residues responsible or determine whether the optimal clone has been identified.

Partial randomisation, incorporating low redundancy codon usage (Balint and Larrick, 1993), is one way of extending the range of a mutagenesis scheme while ensuring that

all possible sequence combinations are cloned. For example, partial randomisation, which introduces eight different amino acids over a stretch of eight residues (assuming one codon per amino acid), will generate $1x10^7$ variants, which can be easily cloned in *E. coli*. Total randomisation of the same eight residues would require an excess of 10^{12} clones (32^8).

In the Gloop2 scFv DNA sequence, four residues were semi-randomised using the codons ag(a/t), (c/t)at, gt(c/t) and (c/g)(c/t)a and assuming an unbiased nucleotide incorporation which will result in 32 different combinations. The transformations resulted in 10^8 colonies and therefore all possible mutants, or combinations of them, were covered with this library size.

Phage prepared from this library were panned three times using biotinylated enkephalin and streptavidin coated beads. After the third round of selection the eluted phage were used in phage ELISA against BSA conjugated enkephalin and BSA conjugated loop peptide. Gloop2 phage was prepared in parallel with the library and used in the same ELISA assays. Serial dilutions of the antigens were used to determine specific binding. Fig.4.6 shows the results from the phage ELISA of the library and Gloop2 phage against BSA-enkephalin and BSA-loop peptide. From these graphs it is easily concluded that the library shows tighter binding to BSA enkephalin than to BSA loop peptide and this is indicated by the shift of the curve about one log to the left. On the other hand the binding of Gloop2 phage shows no significant difference between the two antigens and an apparent non-specific binding to BSA-enkephalin.

There are a number of possible explanations for the non-specific binding of the loop peptide to GlaMor. For example, it may have been due to the existence of multivalency, which would cause even weak binding to show up. To a first approximation, this means that if there are 3 sites each at 20% occupancy the affinity of any trivalent phage particle will appear as though it is one site with 60% occupancy, so that even intrinsically low affinity binding will show up. Alternatively, it may have originated in the non-specific binding of phage particles. However, this would require that the loop peptide has a particularly 'sticky' surface and consequently a similar effect should have been seen with the wild type phage binding to loop peptide.

A more likely explanation is that there are residual recognition motifs for the loop peptide present in the Glamor constructs. The retention of all but two of the H2 residues (for reasons previously explained) and of other CDR residues is a weakness of the design, since at least some residues responsible for the loop peptide specificity may have been retained. At the time of writing, the ability of computer programs for modelling antibody combining sites (eg AbM, developed by the Rees group and marketed by Oxford Molecular plc) is much more effective than at the time of the design, so that more extensive changes to H2 (and H3) could now be made while ensuring retention of the overall groove topography.

PCR screening of eight mutant clones from this library identified four clones (50%) carrying inserts of the right size and after sequencing all four clones shared the same sequence. When this was compared with the Gloop2 wild-type scFv sequence it was found that it carried the set of mutations shown in Fig. 4.8. Interestingly this set of mutations matched with the fourth ranked predicted construct from the design process

(Chapter 3). The sequences of the mutant clones carried one additional mutation in CDRH1, position I34R and a PCR error (probably) mutation at position 89 of the light chain where the designed tryptophan was actually a cysteine. At this stage no further mutagenesis was carried out and the GlaMor clone was used to produce soluble scFv.

5

Cloning of scFv genes in expression vectors

5.1 Expression Vectors

All antibody expression strategies (described in Chapter 6), require the use of appropriately designed expression vectors. The main reason is the stress imposed on *E. coli* by expressing the antibody and therefore the protein expression must be carefully controlled to the right level.

Vectors constructed for the functional expression of antibody fragments carry a regulatable promoter and its cognate repressor gene, efficient ribosomal binding sites, structural genes encoding the corresponding antibody fragments fused to bacterial

signal sequences, a transcription terminator, and a high copy number replication origin.

The vectors described here have been tested in expression studies of the Gloop2 and GlaMor scFv fragments in *E. coli.* All the primer sequences used for cloning are summarised in Table 5.1 at the end of this chapter.

5.1.1 pKK322

The vector map of pKK322 and the multiple cloning site are illustrated in Fig. 5.1.

Subcloning of GlaMor scFv gene in pKK322 vector

GlaMor/fd-tet-DOG1 DNA was prepared as described in Chapter 4 from log phase *E. coli* TG1 cells, infected with phage carrying the GlaMor/fd-tet-DOG1. Plasmid DNA was prepared following the LiCl method alongside the Gloop2/pKK322 DNA (provided by Dr. G.E.Elliott, Fig. 5.2I). Both DNA's were double digested with *PstI* and *XhoI* (reaction 1) and fragments of the appropriate size (GlaMor scFv insert and pKK322 vector fragment) were gel purified and extracted with agarase. The ligations (reaction 2) were set up and the reactions were used to transform *E. coli* TG1 competent cells.

Ten randomly selected colonies were PCR screened using primers 1FOR and XhoMutBACK (Table 5.1) and all showed the presence of the correct size cloned insert (Fig. 5.2II). DNA was prepared from two of these clones and was sequenced to verify the cloning of the mutant sequence before commencing the expression trials.

Reaction 1: Restriction digest of Gloop2/pKK233 and GlaMor/fd-tet-DOG1 DNA

1. The following were mixed in order:

15µl of Buffer 3 (New England Biolabs), approximately 1µg of the GlaMor/fd-tet-DOG1 DNA or Gloop/pKK233, 1µl of *Pst*I (diluted to 10units/µl), 1µl of *Xho*I (diluted to 10units/µl) 1.5µl 100X BSA (10mg/ml) and H_2O to a final volume of 150µl.

2. The mixtures were incubated overnight at 37°C

3. The enzymes were heat inactivated at 65°C for 20 minutes.

4. The fragments of the desired size were purified on an 1% (w/v) LMP agarose gel and extracted with agarase.

5. 1µl of each fragment and digested vector was analysed on an 1% (w/v) agarose gel in order to determine the efficiency of the purification and the final concentrations before setting up the ligations.

Reaction 2: Ligation and transformation of competent *E. coli* cells

1. The following were mixed on ice:

2µl of 10X ligase buffer (containing 10mM ATP), 20ng of digested vector, 30ng of digested PCR fragments, 1µl of diluted T4 DNA ligase (4units/µl, New England Biolabs) and H_2O to a final volume of 20µl.

As a control the insert DNA was omitted (-ve control)

2. The mixtures were incubated at 16°C overnight

3. The enzyme was heat inactivated at 65°C for 10 minutes

4. All 20µl of each reaction was used to transform 50µl of thawed competent XL1-Blue cells

5. The transformed cells were centrifuged, resuspended in approximately 200µl of the supernatant and spread onto LB agar plate containing 100µg/ml Ampicillin, 15µg/ml Tetracyclin and 2% (w/v) Glucose
6. The agar cultures were incubated at 37°C for approximately 15 hours.
7. The transformant colonies were counted and are shown in the table below

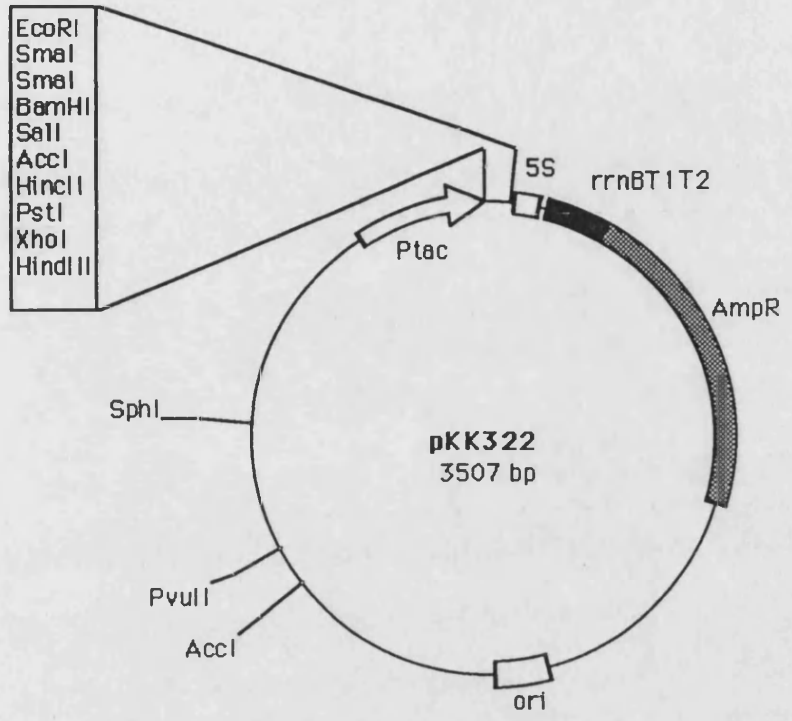
Sample	*DNA Description*	*No of colonies*
1	GlaMor/pKK322	139
2	negative control (no insert)	1
3	positive control (Gloop/pKK322)	>1000

Sequencing of Gloop2 and GlaM positive clones

DNA was prepared for the ten GlaMor/pKK233 positive clones using the QIAprep Spin Plasmid Kit (QIAGEN Ltd.) as described in Materials and Methods.

Two samples were sequenced using the 1FOR and XhoMutBACK primers. Sequences were analyzed on a 377 Automated DNA Sequencer ABI prism (Perkin-Elmer). Sequence analysis was performed using SeqEd (Applied Biosystems) (Appendix I) and verified the successful cloning of GlaMor scFv gene in the pKK322 vector.

I



II

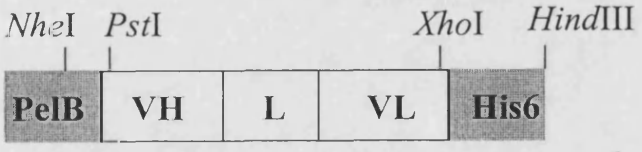


Fig. 5.1: Cloning the scFv genes in the pKK322 expression vector

I. Structure of the pKK322 vector

II. scFv (Gloop2 or GlaMor) cloning cassette in the pKK322 expression vector

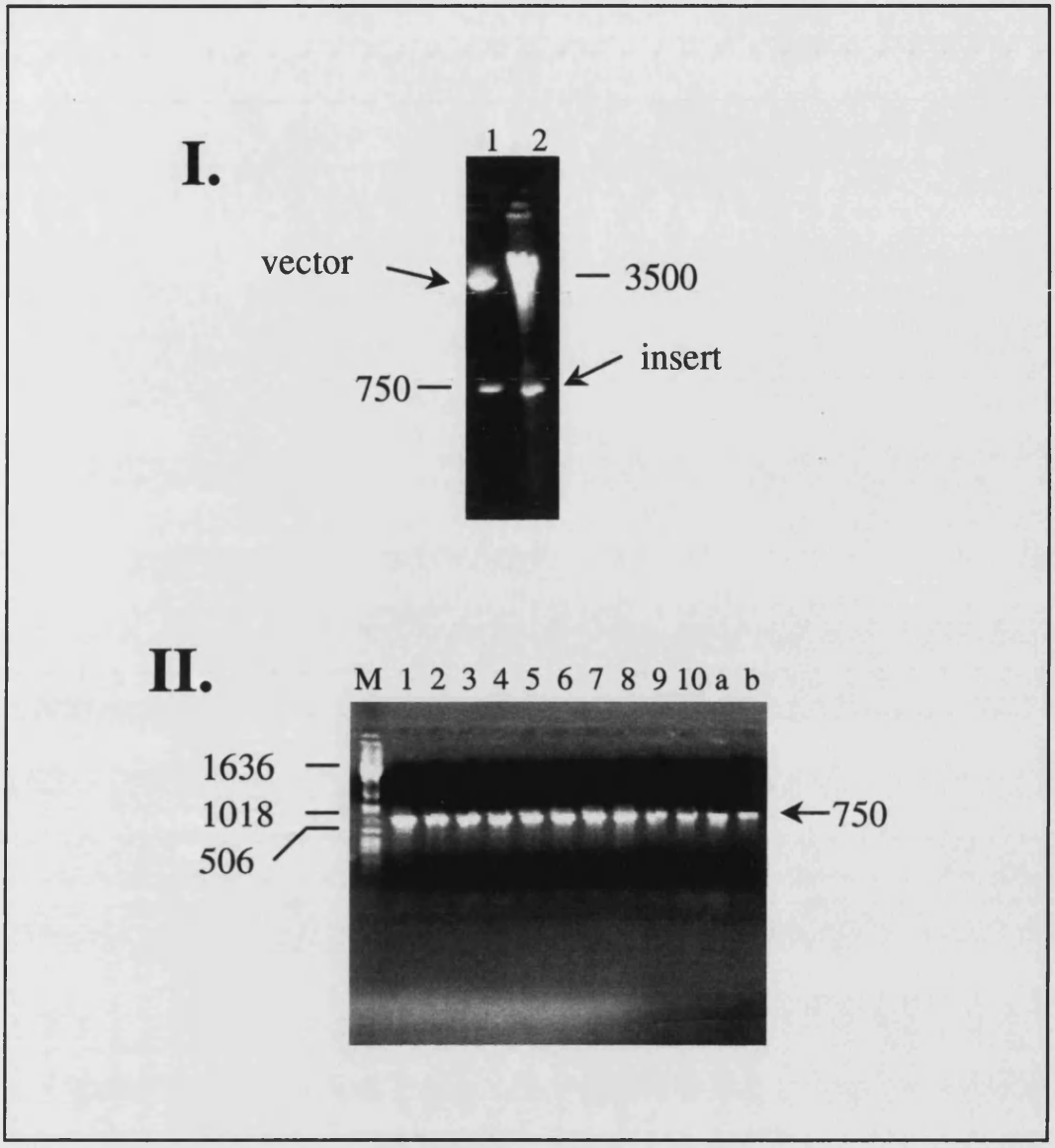


Fig. 5 2: Subcloning of GlaMor scFv gene in the pKK322 vector

I. Double digests of Gloop2/fd-tet-DOG1 (1) and GlaMor fd-tet-DOG1 (2) plasmid DNA with *Pst*I and *Xho*I Both vector and insert DNA were gel purified and the GlaMor scFv gene was cloned as a cassette in the pKK322 vector. Ligated DNA was used to transform XL1-Blue *E. coli* competent cells.

II. Positive clones were identified first by PCR screen of 10 randomly selected colonies (1-10) using 1FOR and XhoMutBACK primers followed by DNA sequencing.

a, +ve control (Gloop/pKK322/TG1); b, +ve control (Gloop/pKK322 plasmid DNA)

All DNA fragments were analysed on 1% (w/v) agarose gels next to 1Kb ladder molecular weight DNA markers (numbers in bp shown on the left).

5.1.2 p*Trc*99A

The p*Trc*99A contains a strong *Trc* promoter which is inducible by IPTG (1-5mM), the lacZ ribosome binding site, and the $lacI^q$ repressor The *Trc* promoter contains the trp (-35) region and the lac UV5 (-10) region separated by 17 base pairs. The lac I^q on the plasmid controls the promoter and makes the plasmid inducible in any *E. coli* host strain. However, this is a very strong promoter and might show very low expression in the uninduced state. The plasmid confers resistance to 100μg/ml ampicillin (Fig. 5.3).

Subcloning Gloop2 and GlaMor scFv genes in p*Trc*99A expression vector.

Gloop2 and GlaMor scFv genes were isolated by digestion with *NheI*/*Hind*III ofGloop2 and GlaMor pKK322 pre-sequenced clones (reaction 1) followed by gel purification of the appropriate size fragments. Because the p*Trc* MCS lacks the *Nhe*I site the wild type vector could not be used for cloning in this case. Instead, the *NheI*/*Hind*III digested vector derived from the digestion of a control scFv fragment pre-cloned in the p*Trc*99A (reaction1) through the *NcoI*/*Hind*III sites followed by gel purification (Fig. 5.4I and II).

The ligations (reaction 2) were set up and were used to transform competent XL1-Blue *E. coli* cells. Recombinant bacteria were screened for the presence of correctly inserted scFv by mini DNA preparations of randomly selected clones followed by restriction digests with *PstI*/*Hind*III. Six clones (three for Gloop2 and three for GlaMor) were screened and all of them showed the presence of a scFv insert of the right size (Fig. 5.4III). DNA of the selected pTrc clones was sequenced using the pelSeqFOR and LS2BACK primers (Table 5.1).

Expression studies using the pTrc vector were interrupted due to the presence of two ATG start codons (one in the *Nco*I site and one in the beginning of the *pel*B leader sequence) upstream of the scFv sequence in different reading frames. Experiments focused on knocking out the first of the two ATG codons (in the *Nco*I site).

Reaction 1: Restriction digest ofGloop2 and GlaMor pKK233 clones and control pTrc vector

1. The following were mixed in order:

5µl of Buffer 2 (New England Biolabs), approximately 500ng eachGloop2 and GlaMor pKK233 DNA or 1µg of the vector with the pre-cloned control scFv, 2µl of *Nhe*I (diluted to 1 unit/µl), 1µl of *Hind*III (diluted to 1 unit/µl), 0.5µl 100X BSA (10mg/ml) and H_2O to a final volume of 50µl.

2. The mixtures were incubated overnight at 37°C
3. The enzymes were heat inactivated at 65°C for 20 minutes.
4. The fragments of the desired size were purified on an 1% (w/v) LMP agarose gel and extracted with agarase.
5. 5µl of each fragment and digested vector was analysed on an 1% agarose gel in order to determine the efficiency of the purification and the final concentrations before setting up the ligations.

Reaction 2: Ligation of Gloop2 and GlaMor *Nhe*I/*Hind*III digested fragments in the pTrc vector and transformation of competent *E. coli* cells

1. The following were mixed on ice: 2µl of 10X ligase buffer (containing 10mM ATP), 10ng of digested vector, 15ng of digested PCR fragments, 1µl of diluted T4 DNA ligase (4 units/µl, New England Biolabs) and H_2O to a final volume of 20µl.

As a control the PCR fragment was omitted (-ve control)

2. The mixtures were incubated at 16°C overnight
3. The enzyme was heat inactivated at 65°C for 10 minutes
4. All 20µl of each reaction was used to transform 50µl of thawed competent XL1-Blue cells
5. The transformed cells were centrifuged, resuspended in approximately 200µl of the supernatant and spread onto LB agar plates containing 100µg/ml Ampicillin, 15µg/ml Tetracycline and 2% (w/v) Glucose
6. The agar cultures were incubated at 37°C for approximately 15 hours.
7. The transformants grown on the plates were counted and are shown in the table below.

Sample	*DNA Description*	*No of colonies*
1	GlaMor/ pTrc99A	220
2	Gloop/pTrc99A	180
3	negative control (no insert)	9
4	positive control (wild type pTrc99A)	>1000

Knocking out the *Nco*I site using Mung Bean Nuclease

Mung Bean Nuclease (New England Biolabs) is a single-strand specific DNA and RNA endonuclease isolated from mung bean sprouts. It degrades single-stranded extension from 3' and 5' ends of DNA and RNA molecules, leaving blunt, ligatable ends. S1 nuclease functions in the same way but has a greater tendency to cleave single stranded DNA across from a nick.

DNA was prepared from previously sequencedGloop2 and GlaMor pTrc clones using the LiCl method and digested with *Nco*I (reaction1, Fig. 5.5I). The digested DNA was ethanol precipitated and treated with Mung Bean Nuclease (reaction 2) for periods of 10 and 30 minutes (Fig. 5.5II). Mung bean treated DNA was ethanol precipitated, added in blunt end ligation reactions (reaction 3) and used to transform competent XL1-Blue *E. coli* cells. DNA was prepared from 10 randomly selected colonies (5 for Gloop2 and 5 for GlaMor) and digested with *Nco*I to screen for the knock out of the ATG codon in the *Nco*I site (*Nco*I^{-ve}) as described in reaction 1 (Fig.5.5III). DNA from the *Nco*I^{-ve} clones was prepared and sequenced across the *pel*B leader sequence using the pelSeqFOR primer.

Reaction 1: Restriction digest of Gloop2 and GlaM pTrc DNA with *Nco*I

1. The following were mixed in order:

   4 µl of Buffer 4 (New England Biolabs), approximately 1µg of Gloop2 or GlaMor p*Trc* DNA, 1µl of *Nco*I (diluted to 3 units/µl),) and H_2O to a final volume of 40 µl.
2. The mixtures were incubated overnight at 37°C
3. The enzyme was heat inactivated at 65°C for 30 minutes.

4. The reactions were ethanol precipitated and the final pellets were resuspended in 5µl of H_2O. 1µl from each reaction was analysed on an 1% (w/v) agarose gel.

Reaction 2: Mung Bean digest of *Nco*I digested DNA fragments

The following were mixed in order:

1. 5 µl of Buffer 2 (New England Biolabs), approximately 400ng of Gloop2 or GlaMor p*Trc* DNA digested with *Nco*I, 1.6 µl of Mung Bean nuclease (diluted to 1 unit/µl),) and H_2O to a final volume of 50µl.
2. The reactions were incubated at 30°C
3. Aftre 10min, 25µl from each reaction was added to 0.01% (w/v) SDS and the rest of the reaction was incubated at 30°C for another 20 minutes before it was inactivated with SDS.
4. All four samples were ethanol precipitated and the final pellet was resuspended in 5µl of H_2O. 1µl from each sample was analysed on an 1% (w/v) agarose gel.

Reaction3: Blunt end ligation of the Mung Bean Nucleased DNA

1. The following were mixed on ice:

3µl of 10X ligase buffer (containing 10mM ATP), 100ng of Mung bean digested DNA (Gloop2 or GlaMor), 1µl of diluted T4 DNA ligase (4units/µl, New England Biolabs) and H_2O to a final volume of 30µl.

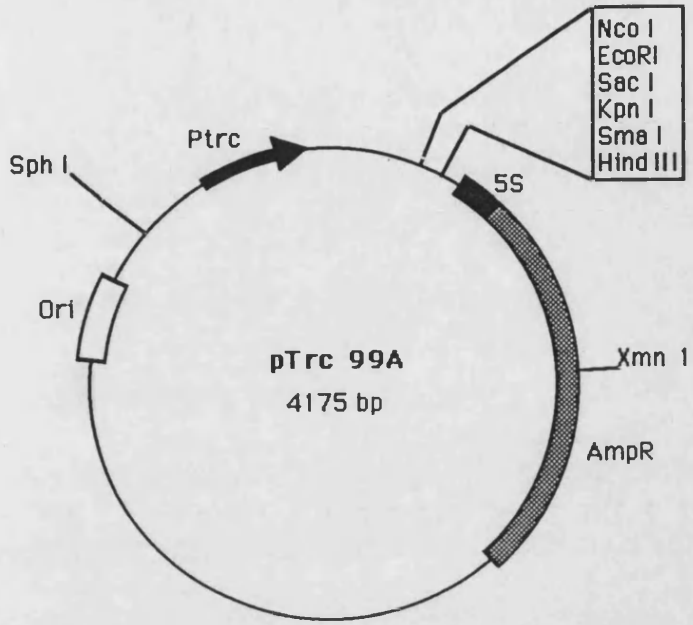
As a control the ligase was omitted (-ve control)

2. The mixtures were incubated at 16°C overnight
3. The enzyme was heat inactivated at 65°C for 10 minutes
4. 15µl of each reaction was used to transform 50µl of thawed competent XL1-Blue cells

5. The transformed cells were centrifuged, resuspended in approximately 200 µl of the supernatant and spread onto LB agar plate containing 100µg/ml Ampicillin, 15µg/ml Tetracyclin and 2% (w/v) Glucose
6. The agar cultures were incubated at 37°C for approximately 15 hours.
7. The transformants grown on the plates were counted and are shown on the table below.

Sample	*DNA Description*	*No of colonies*
1	GlaMor/p*Trc*	24
2	Gloop/p*Trc*	15
3	negative control (no ligase)	0
4	positive control (wild type p*Trc* vector)	>1000

I



II

AGGAAACAGACC ATG GAA TTC GAG CTC GGT ACC CGG GGA TCC

Nco I *Eco*R I *Sac*I *Kpn* I *Sma* I *Bam*H I

TCT AGA GTC GAC CTG CAG GCA TGC AAG CTT

Xba I *Sal* I *Pst* I *Sph* I *Hind* III

III

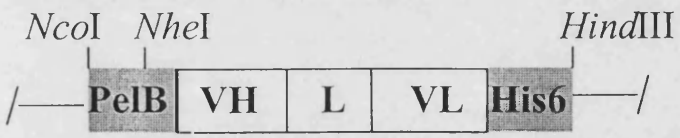


Fig. 5.3 : Cloning the scFv genes in the p*Trc* 99A expression vector

I. Structure of the expression vector p*Trc* 99A

II. Sequence of p*Trc* 99A cloning sites

III. scFv (Gloop2 or GlaMor) cloning cassette in the p*Trc* 99A expression vector

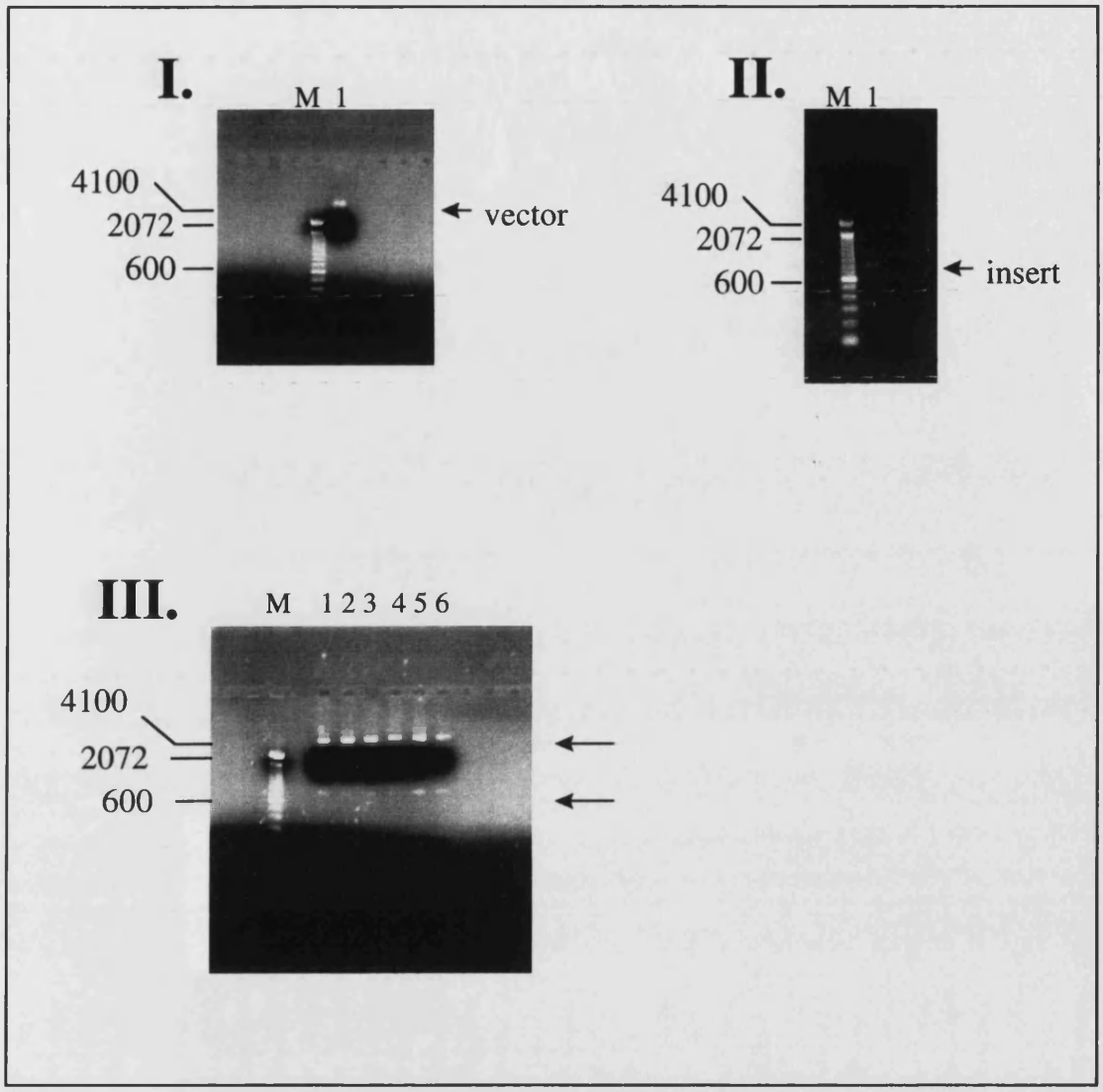


Fig. 5.4: SubcloningGloop2 and GlaMor scFv genes in the p*Trc*99A expression vector

I Vector DNA (1) and II GlaMor scFv insert DNA (1) double digested with *Nhe*I and *Hind*III both after gel purification and before adding to the ligation reactions. The ligations were used to transform XL1-blue *E. coli* competent cells. The same procedure was followed for theGloop2 scFv DNA.

III. Positive clones were identified by plasmid mini-preparations of six randomly selected colonies followed by double digests with *Nhe*I and *Hind*III.

1-3, Gloop2/pTrc and 4-6, GlaMor/pTrc DNA

All DNA fragments were analysed on 1% (w/v) agarose gels next to 1Kb ladder molecular weight DNA markers (numbers in bp shown on the left).

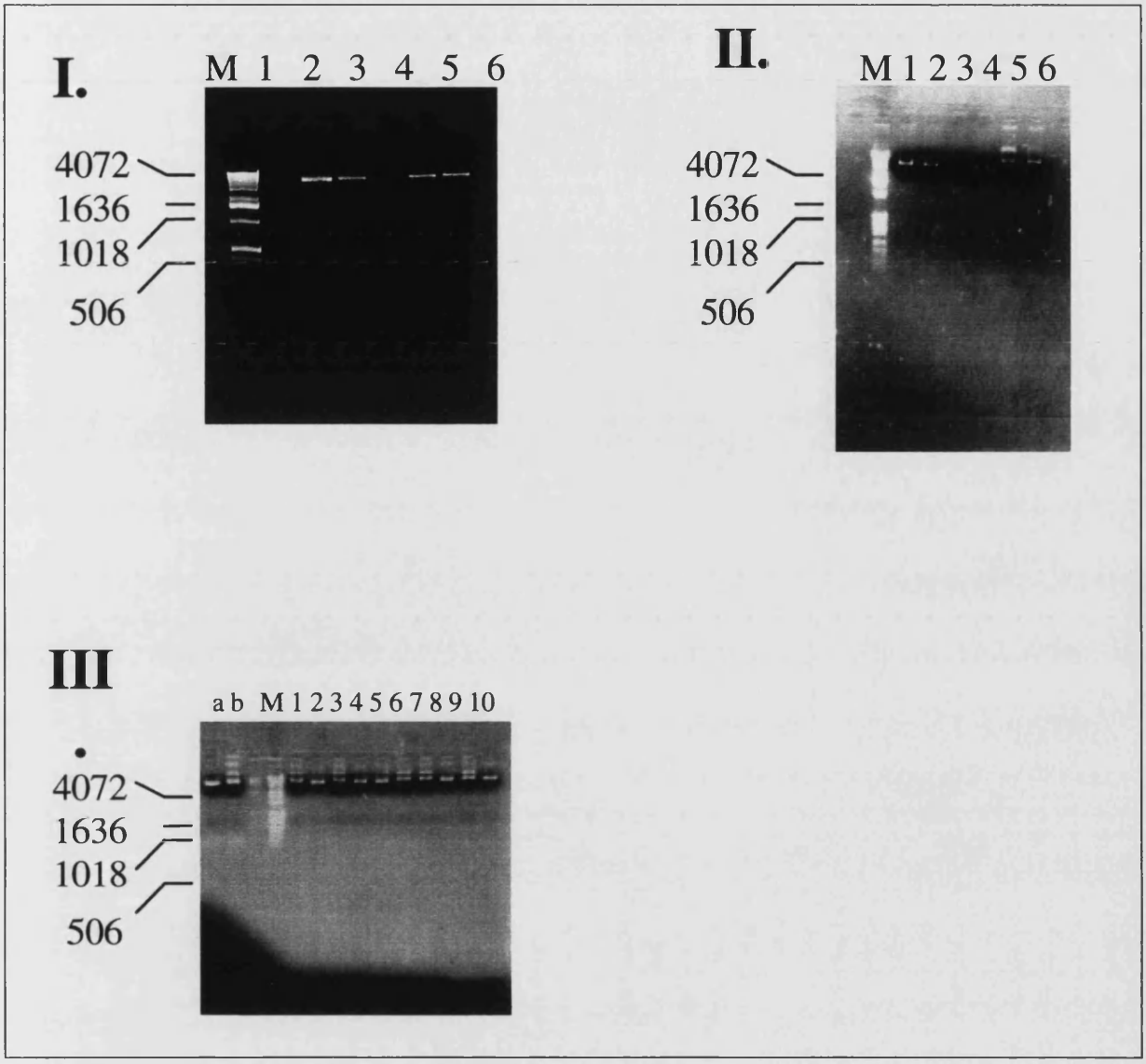


Fig.5.5: Knocking out the ATG codon in the *NcoI* restriction site using Mung Bean Nuclease

I. Gloop2 and GlaMor pTrc DNA was linearised with *NcoI*
1 and 4 (+ve controls), Gloop2 and GlaMor p*Trc* DNA digested with *NcoI* and HindIII
2, 3 and 5, 6, Gloop2/pTrc and GlaMor/p*Trc* respectively digested with *NcoI* (in duplicates)
II. Mung Bean Digest
1 and 3, Gloop2 and GlaMor p*Trc* DNA linearised with *NcoI* and incubated with Mung Bean nuclease at 30°C for 10 minutes; 2 and 4, Gloop2 and GlaMor p*Trc* DNA linearised with *NcoI* and incubated with Mung Bean nuclease at 30°C for 30 minutes
5 (+ve control), Gloop2 p*Trc* DNA incubated with Mung Bean at 30°C for 30 minutes
6 (-ve control), Gloop2 p*Trc* plasmid DNA.
Gloop2 and GlaMor p*Trc* DNA was blunt-end ligated after the Mung Bean treatment and transformed JM109 competent *E. coli* cells. Ten colonies (1-10) were selected randomly, used to prepare DNA and incubated with *NcoI*, to identify positive clones that show deficiency to digest with *NcoI*. Further DNA sequencing was performed to confirm the knocking out of the ATG codon.
1-5, Gloop2 p*Trc* DNA; 6-10, GlaMor p*Trc* DNA; a, (+ve control) Gloop2 p*Trc* DNA digested with *NcoI*; b, Gloop2 p*Trc* plasmid DNA. All DNA fragments were analysed on 1% (w/v) agarose gels next to 1Kb ladder molecular weight DNA markers (numbers in bp shown on the left).

5.1.3 pUC119His6mycXba

This vector is a derivative of pUC119 (Vieira and Messing, 1987) in which the polylinker has been replaced by the sequence shown in Fig.5.6. It has been used successfully for the expression of antibody fragments in *E. coli* and was generously donated by Professor. R. Hawkins (University of Bristol).

This vector was extensively used for all the expression studies of Gloop2 and GlaMor soluble scFv described here.

SubcloningGloop2 and GlaMor scFv genes in pUC119His6mycXba vector

Gloop2 and GlaMor p*Trc* clones were used as template DNA and the scFv genes were PCR amplified (reaction 1) with Xba FOR and Not BACK primers (Table 5.1) that introduced the *Xba*I and *Not*I cloning sites at the 3' and 5' end of the scFv genes respectively (Fig.5.7II). The amplified inserts were column purified, digested with *Xba*I and *Not*I (reaction 2) and gel purified. pUC119His6MycXba vector DNA was prepared using the QIAprep Spin Plasmid Kit (QIAGEN), linearised using the same enzymes (reaction 2) and gel purified.

The ligations were set up (reaction 3) and the reactions were used to transform competent XL1-Blue *E. coli* cells. Recombinant bacteria were screened for the presence of correctly inserted scFv through In-well lysis screening (Epicentre Technologies Co.). Out of the twenty randomly selected colonies, twelve (seven Gloop2 and five GlaMor) turned out to be positive and were used for sequencing prior to expression studies (Fig. 5.7III).

Reaction 1: PCR amplification of Gloop2 and GlaMor scFv fragments

1. The following were mixed in order:

   5µl 10X PCR buffer, 0.5µl 100X BSA (10mg/ml), 5µl of 10X dNTP, 1µl *Xba*I primer (50pM), 1µl *Not*I primer (50pM), 1µl DNA, 1µl Vent DNA polymerase (2u/µl) and 35.5µl H_2O to a final volume of 50µl. The mixture was overlaid with approximately 50µl of mineral oil.

2. The PCR was performed with 30 cycles of 94°C (1min), 55°C (1min) and 72°C (2min)

3. 5µl from each reaction was analysed on an 1% (w/v) agarose gel .

Reaction 2: Restriction digest ofGloop2 and GlaMor PCR products and pUC119 Myc His6 vector

1. The following were mixed in order:

   6µl of Buffer 3 (New England Biolabs), approximately 600ng of the purified PCR products or 1µg of the vector, 2µl of *Xba*I, 2µl of *Not*I, 0.6µl 100X BSA (10mg/ml) and H_2O to a final volume of 60µl.

2. The mixtures were incubated overnight at 37°C

3. The enzymes were heat inactivated at 65°C for 20 minutes.

4. The fragments of the desired size were purified on an 1% (w/v) LMP agarose gel and extracted with agarase.

5. 5µl of each fragment and digested vector was analysed on an 1% (w/v) agarose gel in order to determine the efficiency of the purification and the final concentrations before setting up the ligations.

Reaction 3: Ligation and transformation of competent *E. coli* cells

1. The following were mixed on ice:

   2μl of 10X ligase buffer (containing 10mM ATP), 16ng of digested vector, 32ng of digested PCR fragments, 1μl of diluted T4 DNA ligase (4units/μl, New England Biolabs) and H_2O to a final volume of 20μl.

As a control the PCR fragment was omitted (-ve control)

2. The mixtures were incubated at 16°C overnight
3. The enzyme was heat inactivated at 65°C for 10 minutes
4. All 20μl of each reaction was used to transform 50μl of thawed competent XL1-Blue cells
5. The transformed cells were centrifuged, resuspended in approximately 200μl of the supernatant and spread onto LB agar plate containing 100μg/ml Ampicillin, 15μg/ml Tetracyclin and 2% (w/v) Glucose
6. The agar cultures were incubated at 37°C for approximately 15 hours.
7. The transformants grown on the plates were counted and are shown on the table below.

Sample	*DNA Description*	*No of colonies*
1	GlaMor/pUC119	200
2	Gloop/pUC119	200
3	-ve control (*Xba*I/*Not*I digested vector purified and ligated)	10
4	+ve control (wild type pUC119)	>1000

Sequencing of Gloop2 and GlaMor positive clones

DNA was prepared for all seven Gloop2 and GlaMor positive puc119His6mycXba clones using the QIAprep Spin Plasmid Kit (QIAGEN Ltd.).

Four samples were sequenced (two for Gloop2 and two for GlaMor) using the lacBFOR and 5FOR primers (Table 5.1). Sequences were analyzed on a 377 Automated DNA Sequencer ABI prism (Perkin-Elmer). Sequence analysis was performed using SeqEd (Applied Biosystems, Fig. 5.8).

Name	Sequence	No of bases	Restriction Sites
lacBFOR	agcggataacaatttcacagga	22	
lacFBACK	cgccagggttttcccagtcacga	23	
LS2BACK	accgcttctgcgttctgattt	21	
XbaFOR	tcactagtctagaatgaaatacctattgcctac	33	*Xba*I
NotBACK	ggtggtgggcggccgcccgtttcagctcgagcttggt	37	*Not*I
pelSeqFOR	gttgacaattaatcatccggctcg	24	
XhoMut BACK	ttcggtgctgggaccaagctcgagctgaaacgg	33	*Xho*I

Table 5.1: The primer sequences used for the cloning experiments.
Sequences of the restriction sited that they are introduced are also shown (underlined sequences).

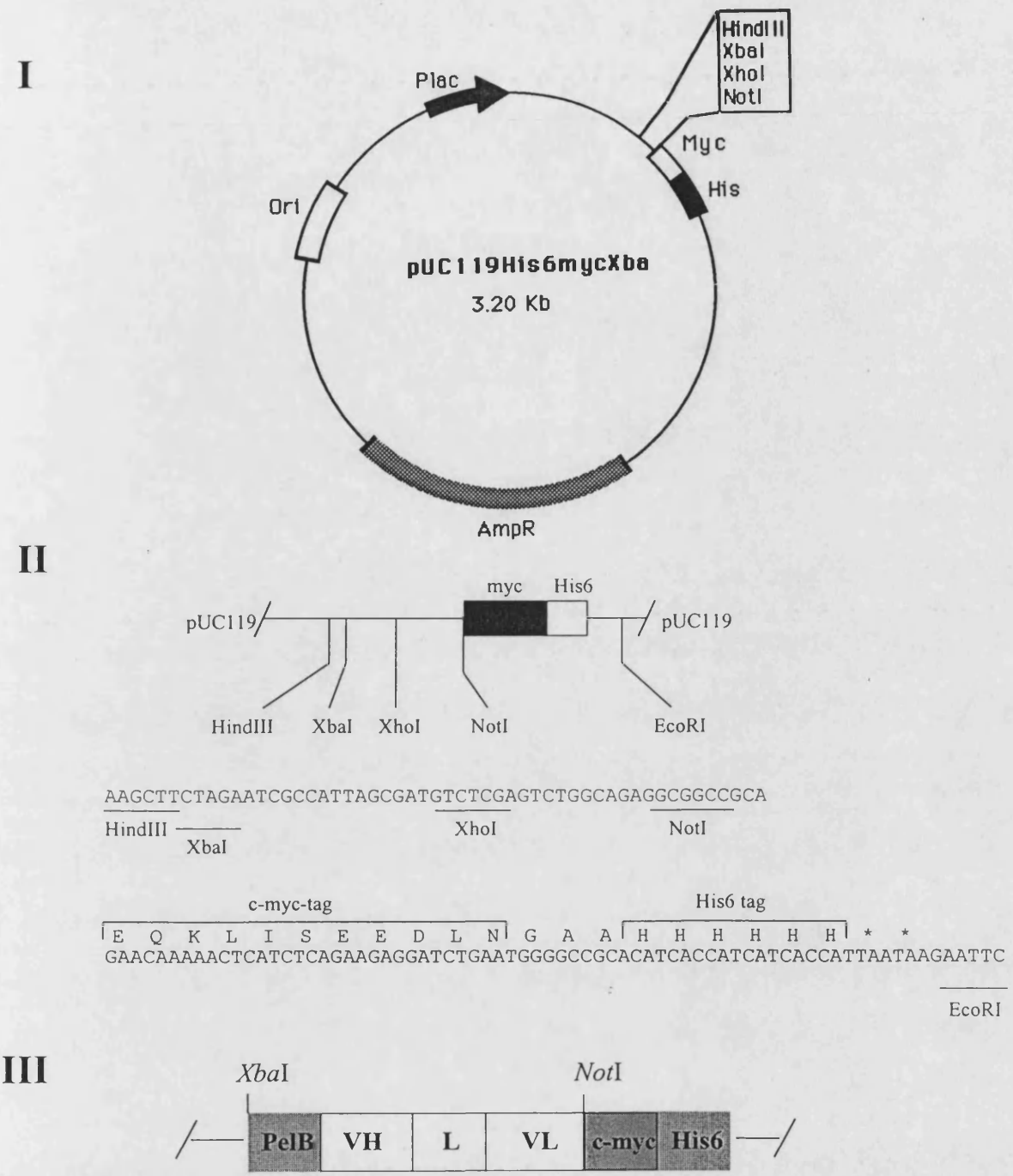


Fig.5.6: Cloning the scFv gene in the pUC119His6mycXba expression vector

I. Structure of vector pUC119His6mycXba

II. Sequence of pUC119His6mycXba cloning site

III. scFv (Gloop2 or GlaMor) cloning cassette in the pUC119His6mycXba vector

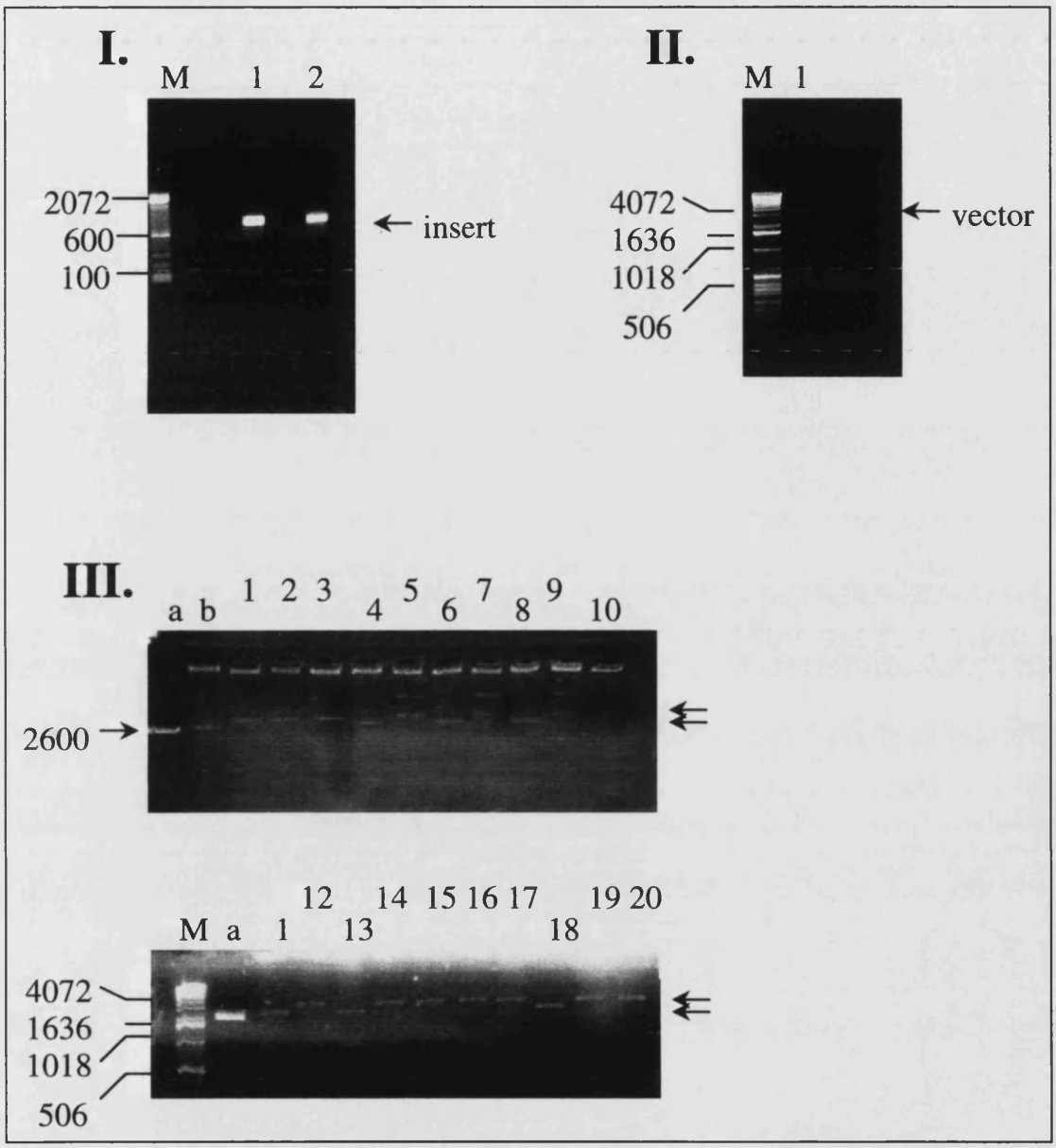
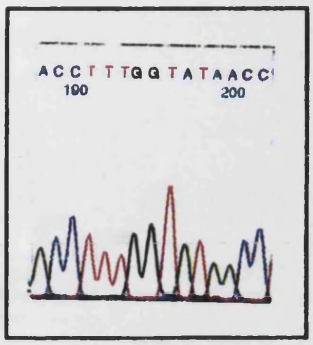


Fig.5.7: SubcloningGloop2 and GlaMor scFv genes in pUC119His6mycXba expression vector
I.Gloop2 and GlaMor scFv genes were PCR amplified using pre-sequenced pTrc clones with the XbaFOR and NotBACK primers.
1,Gloop2 and 2, GlaMor scFv DNA PCR products
II. pUC119His6MycXba vector DNA (1) was prepared from TG1 *E. coli* cells, double digested with *Xba*I and *Not*I, gel purified before it was added to the ligation reaction. Ligated DNA was used to tranform XL-1Blue TG1 *E. coli* competent cells.
III. Twenty colonies were randomly selected and screened by in-well lysis screening. Positive clones were identified by an increase in the M.W of plasmid DNA (top arrows) compared with the M.W of the pUC119His6MycXba wild-type plasmid DNA (lower arrows)
1-10, Gloop2/pUC119His6MycXba; 11-20; GlaMor/pUC119His6MycXba;a, +ve control pUC119His6MycXba plasmid DNA; b, +ve control pUC119His6MycXba/XL-1Blue. All DNA fragments were analysed on 1% agarose gels next to 1Kb ladder molecular weight DNA markers (numbers in bp shown on the left)

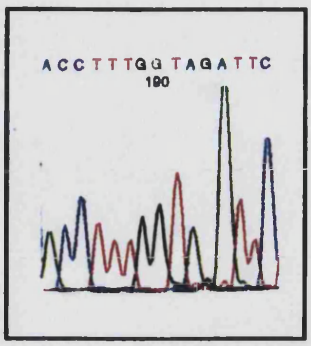
Gloop2 Sequence **Mutation** **GlaMor sequence**

CDRH1

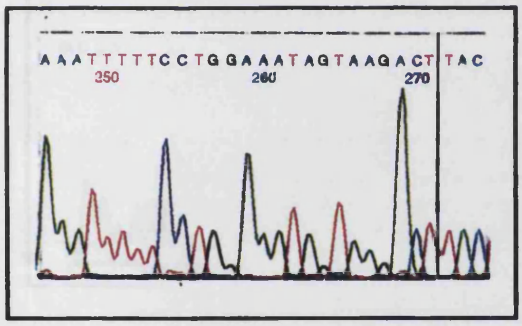


ata *aga*
I 34 R

acc **T 35 F** *ttc*

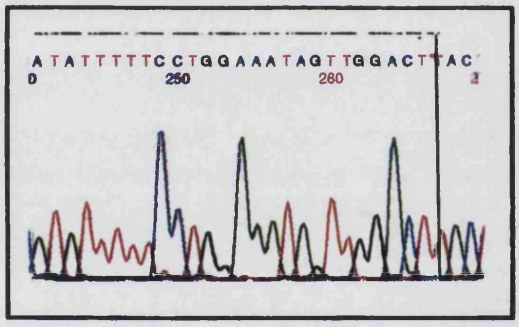


CDRH2

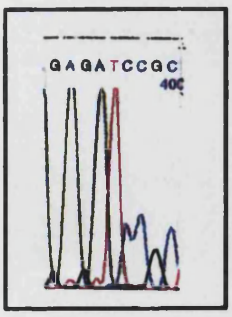


aaa *ata*
E 50 H

aag **K 56 W** *tgg*



CDRH3



gag *cat*
E 95 H

atc **I 96 D** *gac*

cgc *tac*
R 101 Y

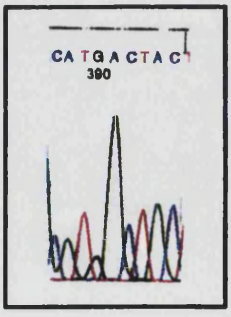


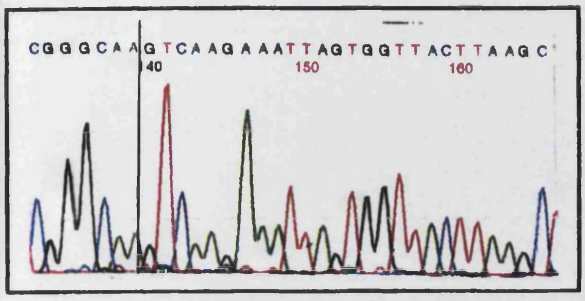
Fig.5.8a: CDR sequences of the VH domain of Gloop2 (left) and GlaMor (right)

The mutations are shown in the middle

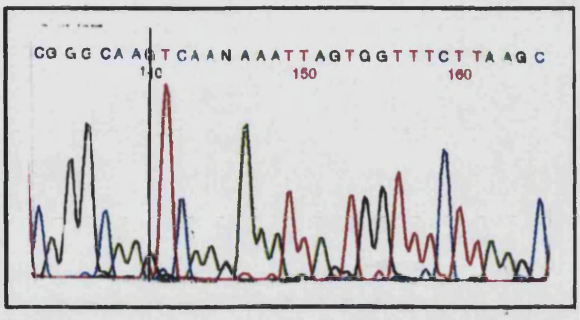
Gloop2 Sequence **GlaMor Sequence**

Mutation

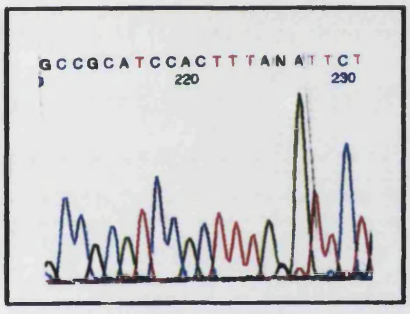
CDRL1

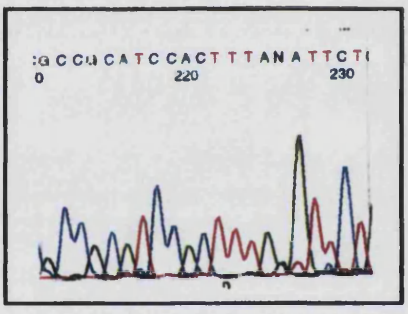


tac **Y 32 F** *ttc*

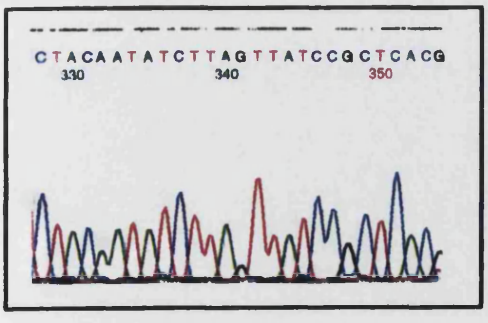


CDRL2





CDRL3



cta **L 89 C** *tgc*

tat **Y 91 D** *gat*

tat **Y 94 F** *ttt*

ctc **L 96 D** *gac*

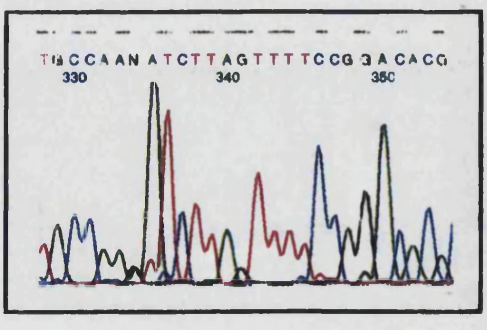


Fig.5.8b: CDR sequences of the VL domain of Gloop2 (left) and GlaMor (right)

The mutations are shown in the middle

5.2 Discussion

Vectors for antibody secretion have been in constant evolution aiming to optimize vector stability, minimize the promoter leak-rate and optimize the expression level in the induced state (Plückthun *et al.*, 1996).

The tight control of the promoter/operator system is one important feature when designing an expression vector. The "leakiness" of expression is responsible for all the unfavorable effects such as periplasmic leakage, plasmid loss, and rearrangement, because of the relatively long time cells spend in pre-cultures, or the main culture before induction. In general the promoter must have a "window" as wide as possible.

For phage display, a strong promoter is of lower importance. In contrast a low background before induction is important since fusion proteins of antibodies with gene3 impose enormous stress on the cell. Soluble expression and phage display have different optima, the latter requiring much stronger transcription and translation. Thus the use of phage-display vectors for soluble expression can be used for characterization but not for efficient production (Plückthun *et al.*, 1996).

Particularly important are tag sequences to follow and purify the antibody conveniently. Purification can be carried out with the histidine tag in combination with metal affinity chromatography (IMAC) (see chapter 6) and the antibody fragment can be detected by a second tag, such as the myc tag (Evans, 1985).

Outside the expression cassette, the vector must code for an antibiotic resistance. In phage display, the expression of the antibiotic resistance gene must not be too high, or low phage titre will be observed, perhaps because of interference between the single-strand production and transcription (Spada *et al.*, 1997).

The scFv cloning cassettes and the different vectors used here are summarised in Fig. 5.9. Attempting to express soluble GlaMor, the scFv gene was initially cloned from the fd-tet-DOG1 phage vector (Hoogenboom *et al.*, 1991) into the pKK322 expression vector. This was facilitated by the Gloop2 scFv pre-cloned in this vector.

As described in the following chapter, the expression studies using the pKK322 vector resulted in very low yields of soluble scFv although several expression experiments were set up and different conditions were tested.

The p*Trc*99A vector was the second vector (Amann *et al.*, 1988), tested due to the very strong trc promoter that would minimize the risk of leakiness prior to induction. During expression experiments this vector was found to be interacting with the host cells resulting in background expression of a 30kD non-specific protein after induction. Additionally, the p*Trc*99A vector contained an extra ATG start codon in the *Nco*I cloning site, upstream of the *pel*B leader sequence. Due to the fact that the second ATG start codon at the beginning of the *pel*B was in a different reading frame to the first ATG, the transcription of the scFv gene could not proceed normally.

The Mung bean Nuclease, used to knock out the first ATG codon, seemed to digest single stranded DNA at a rate difficult to control, even though the reactions were carried out over a range of time intervals.

The last vector tried was the pUC119HismycXba (Griffiths *et al.*, 1994) which was provided by Prof. R. Hawkins and had been used successfully for the expression of soluble antibody fragments (Griffiths *et al.*, 1994). The additional advantage of this vector was the existence of the *c-myc* tag sequence upstream of the His tag that allowed detection of scFv expression using the anti-myc tag antibody, 9E10 (Evans, 1985).

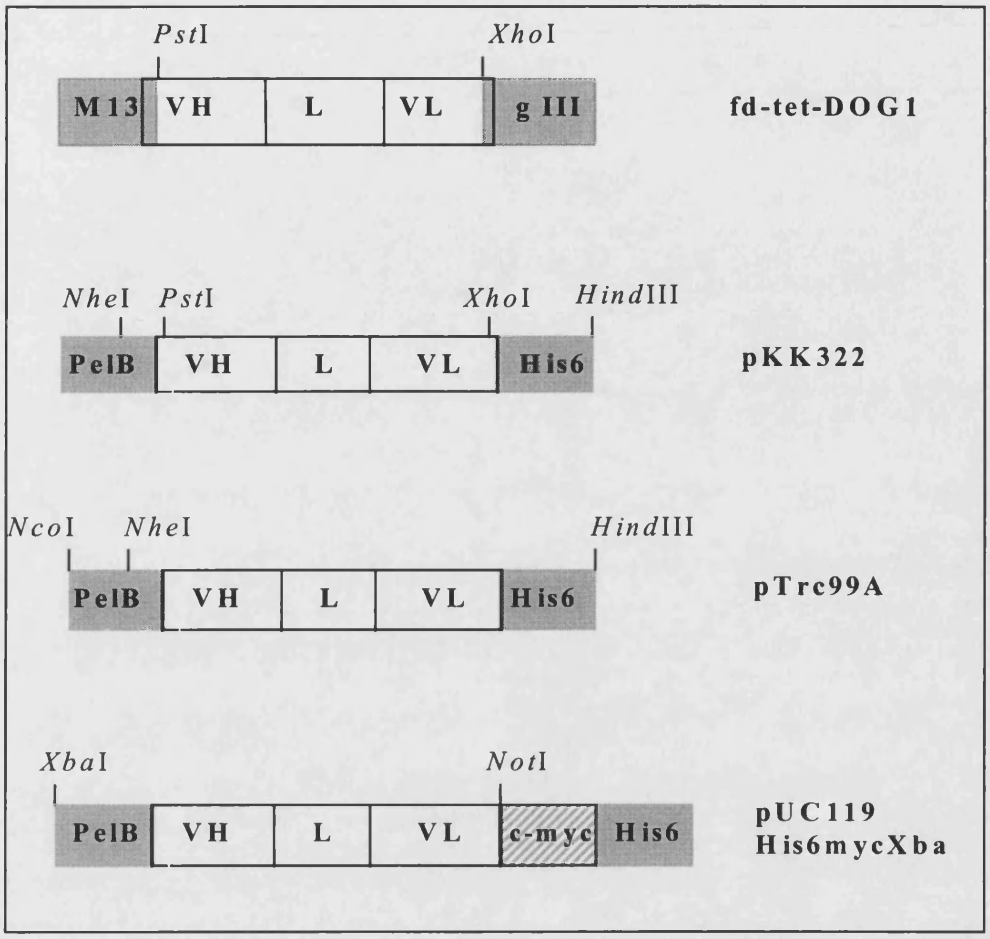


Fig. 5.9: Schematic drawing of the scFv cloning cassettes in different vectors.

6

Bacterial expression and purification of soluble scFv

6.1 Introduction

6.1.1 Expression Strategies

Two basic strategies can be employed to obtain antibody fragments from *E. coli*. The first is the functional expression of correctly folded fragments by secretion into the periplasmic space, (Skerra *et al.*, 1991) and the second is the *in vitro* refolding of protein obtained from inclusion bodies, either from the cytoplasm or the periplasm. There are several advantages to this secretory expression system. First it directly leads to an assembled functional product with correctly formed disulphide bonds without the need to refold the protein *in vitro* (Glockshuber *et al.*, 1992) Second the problem of protease degradation is greatly diminished, as there are fewer proteases in the periplasm than in the cytoplasm. However, there are studies where soluble functional scFv expressed in the cytoplasm was about five times higher that that in the periplasmic extracts from identical numbers of *E. coli* cells (He *et al.*, 1995).

The periplasmic space is a subcellular compartment located outside the cytoplasmic membrane, but retained within an outer cell wall which is impermeable to molecules greater than 600-800kDa (Nikaido and Rosenberg, 1981). Bacterial leaders are employed to direct the protein beyond the cytoplasmic membrane into the periplasmic space which may serve as a functional equivalent to the eukaryotic endoplasmic reticulum (Skerra and Plückthun, 1988a). This is to be distinguished from the proteins that are destined for export beyond the outer cell wall and require specific gene systems the least complex of which needs three separate genes (David, 1990). A variety of bacterial signal sequences have been shown to direct transport and correct cleavage (Plückthun, 1991). The signal peptide sequence, bacterial *pe*ctate *l*yase (*pelB* leader) from *Erwinia carotovora* (Lei *et al.*, 1987) has been fused at the amino termini of antibody fragments and facilitated their transport through the inner bacterial membrane (Dubel *et al.*, 1993). Somerville and coworkers (1994) found that scFv directed against human carcinomas was not folded correctly in the absence of a bacterial leader segment. It is often reported that active material can be found in the culture supernatant of *E. coli* expressing an Ig fragment with a bacterial leader segment (Better, M. *et al.*, 1988). This has been mistakenly referred to as bacterial secretion but careful analysis reveals that appearance of the heterologous protein in the supernatant coincides precisely with cell death and lysis (Takkinen *et al.*, 1991). Recently, Kirpriyanov *et al.* (1997a) found that the scFv could be made to accumulate in the periplasm or secreted into the medium by simply changing the incubation conditions and the concentration of the inducer.

6.1.2 Transcriptional regulation of the *lac* promoter

Different expression vectors (see chapter 5) with a variety of promoters have been used to direct transcription of antibody fragments in *E. coli*. The transcription of the foreign gene is repressed to avoid a potential lethal effect on the cells until the appropriate time, then induced at the optimal level. However, upon induction, it appears that the level of expression is difficult to regulate. The *lac* promoter from the *E. coli* lactose operon is one of the promoters most commonly used to regulate the expression of recombinant genes in bacteria (Auger and Bennet, 1987). The *lac* promoter can be induced by lactose or iso-propyl β-D-thiogalactopyranoside (IPTG), and thus, incremental regulation of transcription can be effected by graded concentration of the inducer. This promoter is also subject to repression by the *lac* I gene product (Beckwith, 1987). However, since a chemical equilibrium exists between bound and unbound repressor molecules, the operator site is not continuously occupied by the *lac* repressor and thus there is generally a low basal level of transcription of the *lac* genes. Trancription from the *lac* promoter is regulated by the binding of the catabolite activator (CAP) to the promoter region (Beckwith, 1987). When CAP binds to the promoter, it increases the affinity of the promoter for RNA polymerase, thereby increasing transcription of the *lac* genes. The affinity of CAP for the promoter is enhanced by its association with cyclic adenosine monophosphate (cAMP). Cellular levels of cAMP are higher when the amount of glucose in the medium is lowest (Fig. 6.1).

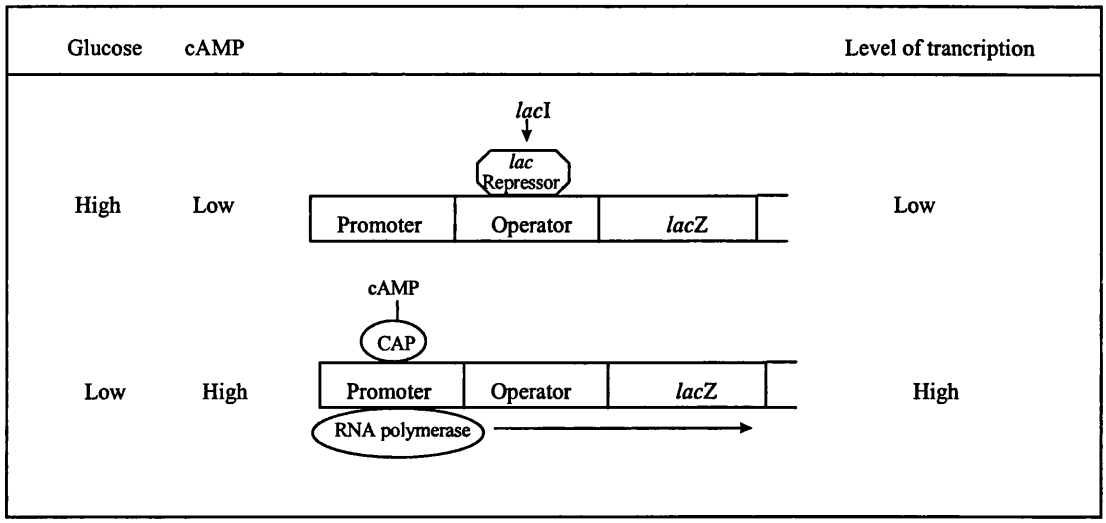


Fig. 6.1: The effect of glucose concentration in the growth medium on the level of transcription from the *lac* promoter in *E. coli*

6.1.3 Factors that influence the expression of antibody fragments in *E. coli*

The amount of IPTG concentration, temperature, composition of the growth medium, the duration of the induction phase and the host cell strain, are factors that influence the expression of antibody fragments in *E. coli.*

Induction temperature and IPTG concentration

Dübel *et al.* (1993) studied the effect of different induction temperatures and IPTG concentrations on the production of scFv in *E.coli* and concluded that concentrations above 100μM IPTG or induction at temperatures above 30°C, led to substantially decreased growth rate and finally to cell death. They suggested that the amount of processed expressed protein was limited by the capacity of the membrane transport system. While the maximum specific growth rate for *E. coli* occurs at a temperature of 37-39°C (Ingraham, 1987) and this coincides with the maximal activity of the *lac* promoter (Auger Benett, 1987) the use of suboptimal growth temperatures in some cases can reduce unwanted metabolic responses to the synthesis of a foreign protein. As a consequence, this improves the yield and/or solubility of the target protein product. For example, total and functional yields of Fv and Fab fragments, were consistently greater at 20-30°C compared to those at 37°C (Skerra *et al.*, 1991). In other studies, functional yields of scFv and Fab fragments were also substantially improved by reducing the culture temperature (Somerville *et al.*, 1994 and Takkinen *et al.*, 1991).

Induction duration

A number of studies have been reported that secreted foreign proteins can be isolated from the culture medium (Better *et al.*, 1988, Skerra and Pluckthun, 1988, and Takkinen *et al.*, 1991). While *E. coli* does not normally secrete proteins to the medium, increasing the duration of the induction phase enhances the release of periplasmic proteins to the surrounding environment. For example, periplasmic Fab fragments leaked into the culture media as the induction phase was extended beyond 10 hours (Shibui and Nagahari, 1992). For the production of a scFv fragment, 10% of the target protein was found in the culture medium after 4 hours, while the proportion of extracellular product increased to 40% after 8 hours and to 90% following 20 hours of induction (Takkinen *et al.*, 1991). In other studies, long induction periods (16-24 hours) have been used to produce 2-10 mg/L yields of antibody fragments in the culture medium (Better *et al.*, 1988 and Takkinen *et al.*, 1991). While in some cases the release of recombinant proteins to the medium has been associated with cell lysis (Somerville *et al.*, 1994), periplasmic leakage has also been associated with actively growing cells (Chalmers *et al.*, 1990). Knappik and Plückthun (1995) studied the optimal induction period of three mutated Fv fragments of the phosphocholine antibody AbMCP603 (Sato *et al.*, 1986) and found that it was dependent on cell stability during the induction periods that were varied between 3, 5 and 12 hours.

The release of periplasmic contents was unlikely to be due to *de novo* synthesis since the total amount of β-lactamase remained at a steady level while release was occurring, as reported by Sommerville and coworkers (1994).

Composition of the growth medium

Studies examining the expression of recombinant proteins under the trancriptional control of the *lac* promoter have shown that the composition of the growth medium at or during the induction phase can significantly affect foreign protein expression (Sambrook *et al.*, 1989). Providing additional amino acids by supplementing the medium with casamino acids, peptone or yeast extract during induction has been shown to improve foreign protein expression (Nancib *et al.*, 1991) and stability (Whitney *et al.*, 1989). In addition, enhancing cell yields, by supplementing complex LB medium with a 0.5-1.5% (w/v) yeast extract during induction improved specific and/or overall recombinant protein expression in some strains of *E. coli* (Li *et al.*, 1990). The addition of sucrose to the growth media was demonstrated to inhibit the aggregation of secreted scFv (Kipriyanov *et al.*, 1997a).

The effect of host cell strain

Different *E. coli* strains, including HB101, JM101, JM83, MC4100, MC1061 and DH5-α, have been assessed for their effectiveness in antibody expression studies (Darveau *et al.*, 1992). TG1 (Anand *et al.*, 1991), XL1-Blue (Kipriyanov *et al.*, 1997a) and TOPP2 (Ulrich, H., oral communication) are popular for the expression of recombinant antibody fragments *in E. coli* strains. Limited leakage or lysis before and after induction and low protease levels are features that characterise those *E.coli* strains useful for expression studies. Plückthun et al (1995) have used successfully the JM83 strain (a K12 strain) and claim that it performs better than other strains. However, they could not suggest a genetic basis for this relative robustness or find any obvious correlation with the genetic markers indicated.

The folding problem

To evaluate different expression systems, two points must be kept in mind. First, the only relevant quantity is the amount of correctly folded, purified protein obtainable at the end. Second, because of the individuality of the antibody variable domain primary sequences and their dramatic influences on the yield of *in vivo* and *in vitro* folding, a comparison of two expression systems with two different antibody fragments may be misleading. Folding *in vivo* can be a limiting process in the production of many recombinant proteins. The individuality of the primary sequence of the antibody variable domain and its importance on folding and subsequently on the expression yields was demonstrated by Knappik and Pluckthun in 1995, when different antibody fragments were expressed under identical conditions. Cell lysis, usually caused by the onset of antibody expression, was diminished by simple point mutations in the antibody protein sequence and this suggested that the most effective strategy to obtain high-yield folding of periplasmic proteins was to engineer the protein itself (Knappik and Pluckthun, 1995). In the same study, they proposed that the crucial determinants are loops in the structure of β–barrel proteins and they described point mutations in these regions that appeared to inhibit aggregation and stabilize the interaction of these regions with the rest of the protein.

Kipriyanov *et al.* (1997b) identified two amino acid residues that were critical for the high level production of scFv in *E. coli* without affecting the antigen binding. It is possible that these changes influence the correct folding of the VH domain via a folding intermediate, but not the affinity.

The scFv antibody fragments are usually correctly processed in the periplasm, contain intramolecular disulphide bonds and are soluble (Glockshuber, 1990). Focusing on

disulphide formation, Proba *et al.*, (1997) observed that under some refolding conditions an unpaired cysteine (when there is no other cysteine to form S-S bond) in the VH3 domain might form a wrong disulphide and result in a misfolded, aggregation-prone product. Interestingly, in the same report when the missing cysteine was derivatized by glutathione, expression levels of functional scFv increased to 50mg/L. Experiments using air oxidation instead of glutathione produced 2,5 mg/L functional scFv fragment. In different studies, Samuelsson *et al.* (1996) described that the requirement for proteins with disulphide bonds to break and reform these bonds, might account for the need to have a low molecular weight redox buffer present during expression in *E. coli.*

6.2 Periplasmic Expression of soluble scFv

A general protocol is presented here, followed in most of the expression experiments described. Departures from this protocol will be indicated where they occur.

Protocol:

1. TG1 or TOPP2 cells freshly transformed with Gloop2 or GlaMor /pUC119His6MycXba respectively were transferred with a toothpick from a plate in 750ml XLM media (16gr yeast extract, 16gr tryptone and 10gr NaCl) containing the appropriate antibiotics (100μg/ml Ampicillin for the TG1 cells and 100μg/ml Ampicillin and 15μg/ml Tetracyclin for the TOPP2 cells) and 0.1% (w/v) Glucose. Glucose helps suppress the *lac* promoter when strains were propagated for plasmid manipulation or long term storage.
2. Cells were grown at 30°C for 24hs shaking at 250rpm to A_{600}nm=4.0 before induction.
3. IPTG was added to a final concentration of 1mM and the induction continued for 3hs at 26°C shaking
4. Cells were harvested by centrifugation at 5,000rpm, for 10min at 4°C
5. Pellets were resuspended in 150ml periplasmic lysis buffer (30mMTris HCl, 20% (w/v) sucrose pH 8.0)
6. 1mM EDTA was added and the lysate was incubated at room temperature for 10min shaking
7. The lysate was centrifuged at 7,000rpm for 10min at 4°C
8. To complex the EDTA from the previous step, the pellet was resuspended in 150ml $MgSO_4$ 5mM and incubated at 4°C for 10min, shaking

9. The lysate was centrifuged again at 7,000rpm for 10min at 4°C
10. The supernatant ("periplasmic fraction") was passed through a filter of pore size 0.45μm, under vacuum before purification

6.3 Immobilised metal affinity chromatography (IMAC) of soluble scFv fragments

In the periplasmic secretion system, considerable enrichment of the antibody protein is possible by cell fractionation, which can be carried out on a preparative scale. The periplasm of *E. coli* contains only a small part of the total cell protein, and therefore subsequent purification is simplified.

A scFv fragment can be purified to homogeneity in a single step using IMAC, and only a minimal structural perturbation of three additional C-terminal residues is necessary. This affinity tail does not disturb the binding of the antibody fragment or interfere with secretion in *E. coli* (Fig. 6.2).

Skerra *et al.*(1991) tested several alternatives to the histidine tail and concluded that five or six consecutive histidine residues give the best binding and expression properties, with minimal structural perturbation.

Deonarain *et al.* (1997) reported that although six histidine residues at the c-terminus of the scFv fragment were used successfully for large-scale clinical production of antibodies this was not sufficient to separate it properly from all *E. coli* proteins.

Protocol

The scFv was purified from the "periplasmic fraction" by IMAC (Hochuli *et al.*, 1988) as follows:

1. 500µl of Ni-NTA resin (QIAGEN) were added to 100ml of periplasmic lysate and were incubated at 4°C under moderate shaking in a glass container, for 30min.
2. The beads were collected by centrifugation at 1,000rpm for 1min, at 4°C at the bottom of a centrifuge tube.
3. The beads were transferred in a home-made column and were allowed to set for 30min.
4. The column was washed extensively with wash buffer (50mM NaH_2PO_4 pH 8.0, 300mM NaCl, 20mM Imidazole) and the flow-through was measured spectrophotometrically at 280nm. The washed were carried out until the OD_{280} of the flow-through was equal to this of the wash buffer (approximately 30-40x of the bed volume).
5. The bound scFv was eluted by applying 10ml elution buffer (50mM NaH_2PO_4 pH 8.0, 300mM NaCl, 250mM Imidazole) and collected in 1ml fractions.
6. The fractions were measured spectrophotometrically at 280nm and fractions corresponding to A_{280} peaks were pooled, concentrated using centicon-10 microconcentrators, buffer exchanged with PBS to a final volume of 200µl and analyzed by SDS-PAGE electrophoresis.

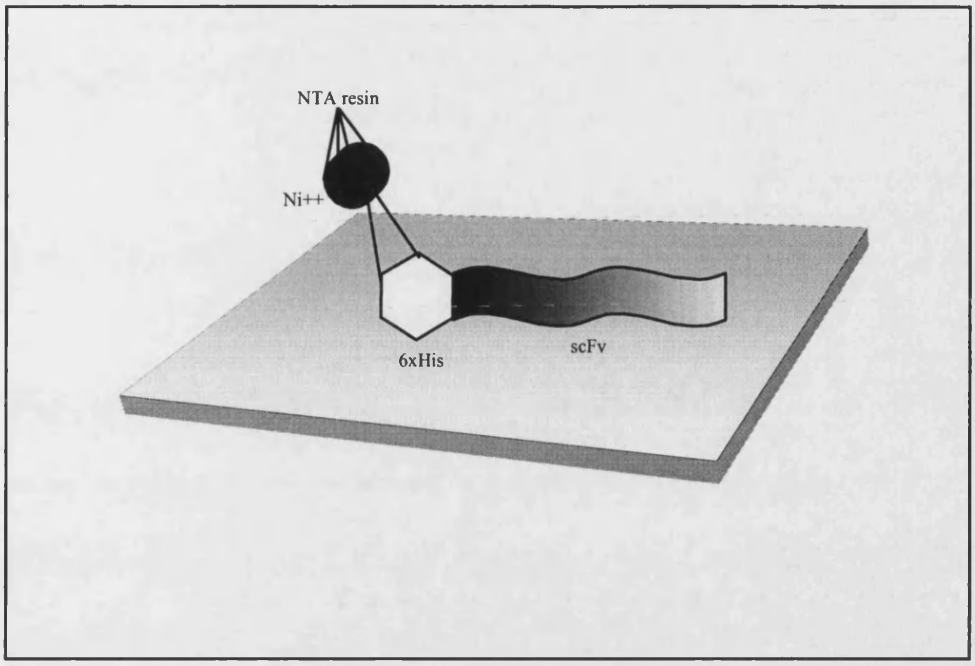


Fig. 6.2: The IMAC mechanism using the Ni^{++} NTA resin (QIAGEN™)
6xHis-tagged scFv from crude lysate of *E. coli* cells binds efficienlty to Ni-NTA agarose. The resin can afterwards be collected and used for conventional column purification.

6.4 Analysis of Gloop2 and GlaMor soluble scFv fragments.

6.4.1 SDS-PAGE Electrophoresis

15µl from each sample was analysed on a 12% SDS-PAGE gel. The gel was stained with Coomassie brilliant blue, dried under vacuum as described in Materials and Methods and retained (fig. 6.2I).

6.4.2 Western Blot Analysis

For the western blot analysis, Gloop2 and GlaMor purified scFv's were fractionated by electrophoresis and blotted onto cellulose. Detection was performed with 9E10 (Evan *et al.*, 1985) which binds the c-myc epitope as the primary antibody followed by a peroxidase- conjugated goat anti-mouse IgG as the second antibody.

Control experiments included *E. coli* cells carrying the wild-type pUC119HisMycXba plasmid (negative control, -ve) and cells transformed with a pre-tested Fab fragment cloned in pUC119HisMycXba (positive control, +ve) (Fig. 6.3II).

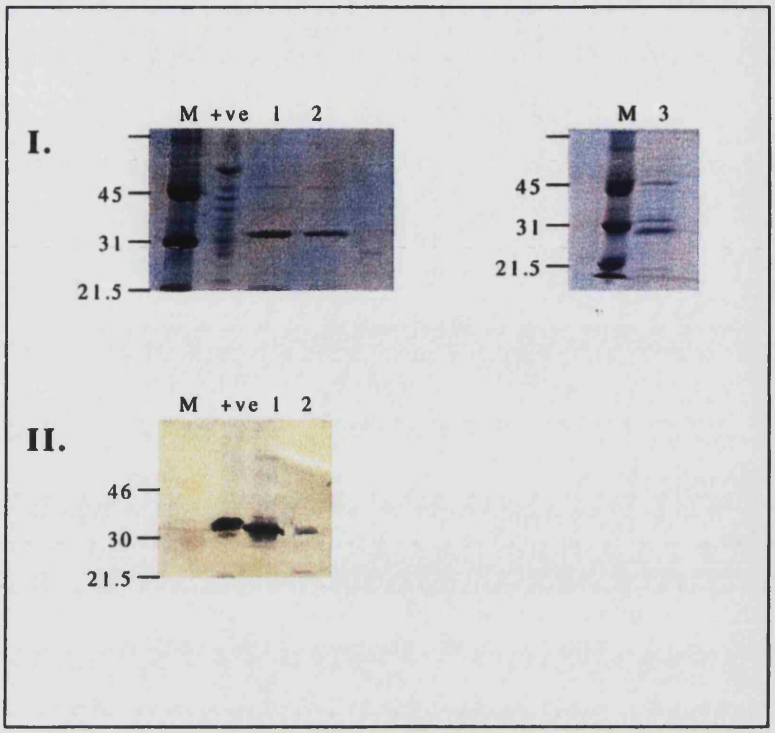


Fig. 6.3: Reducing 12% SDS-PAGE (I.) and immunoblot analysis (II.) of the periplasmic expression of Gloop2 and GlaMor scFv.

Periplasmic lysates were subjected to IMAC purification and selected fractions were concentrated and buffer exchanged in PBS. 15µl from each sample was analyzed per lane.

I. Proteins were stained with Coomassie Brilliant Blue

1 and 2 two different fractions of purified Gloop2 scFv, 3; purified GlaMor scFv; +ve control; unpurified periplasmic scFv.

II. Western Blot Analysis was performed with 9E10 Ab as a primary and a peroxidase-conjugated goat anti-mouse polyclonal as a second Ab.

1; Gloop2 scFv, 2; GlaMor scFv, +ve; control scFv

The molecular weight markers (M) shown on the left have protein bands of relative molecular weights of 21.5; 31;and 45 kDa for the SDS-PAGE analysis, and 21.5; 30; and 46kDa for the immunoblotting.

6.5 Optimisation studies of Gloop2 and GlaMor scFv expression.

6.5.1 The effect of induction duration

The effect of the induction duration on the expression was studied for the Gloop2 and GlaMor scFv's in the pUC119His6MycXba system. Experiments were performed for 24 hours induction periods and the culture densities were recorded by measurements of the absorbance (A) at a wavelenght of 660nm. Both the periplasmic and the supernatant fractions were analysed for the appearance of a correctly folded scFv fragment by SDS-PAGE and Westen Immunoblotting after the 24 hour induction period (Fig. 6.4).

After 24hrs incubation, Gloop2 and GlaMor scFv fragments were detected only in the culture supernatant. No detectable amounts of scFv could be found in the periplamic fractions.

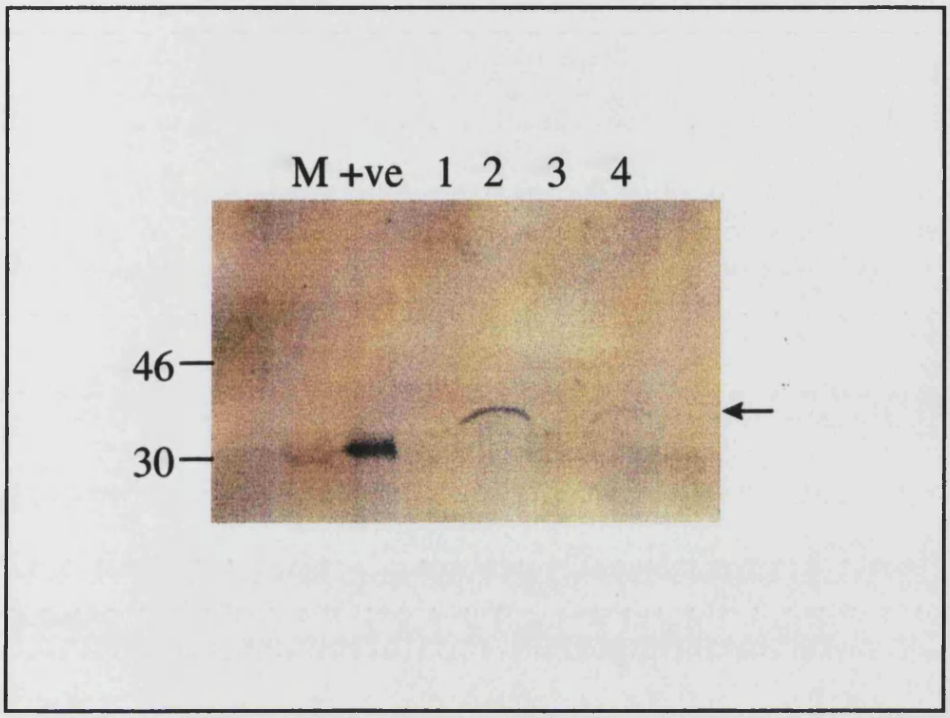


Fig. 6.4: Immunoblot analysis of the periplasmic fraction and culture supernatants from *E. coli* cells expressing Gloop2 and GlaMor scFv.

Periplasmic lysates and culture supernatants were analysed after 24 hours of induction. Cells were harvested, lysed and 4ml from each sample was concentrated and analysed on a reducing 12% SDS-PAGE gel followed by Western Blot analysis.

1 and 2; Gloop2 2, 3; and 4 GlaMor scFv of the periplasmic fraction and culture supernatant respectively.

+ve control is an FabRef fragment periplasmic fraction after 3 hours induction.

The molecular weight markers (M) have protein bands of relative molecular weights in kDa shown on the left.

6.5.2 The effect of host cell strain

TG1 and TOPP2 (Stratagene) strains of *E. coli* were used comparatively in the expression studies of each Gloop2 and GlaMor scFv in the pUC119His6MycXba system. The TG1 strain has been used successfully for the expression of antibody

fragments in *E. coli* (Griffiths *et al.,* 1994) while the TOPP2 strain is recommended by Stratagene as a non-K-12 strain suitable for the expression of proteins that are otherwise difficult or impossible to produce in *E. coli* K-12 strain. Gloop2 scFv expression levels were higher in TG1 cells while GlaMor yields were higher in TOPP2 *E. coli* cells.

6.5.3 Expression using a redox buffer system

It seemed worth investigating whether the low expression levels of GlaMor scFv could be attributed to the one extra cysteine (five in total) in CDRL3. The wild-type sequence of Gloop2 scFv that contained the normal complement of four cysteines gave significantly higher expression yields. Expression of GlaMor scFv was attempted in the presence of a glutathione redox buffer in the culture medium during induction.

Protocol:

1. 500ml of XL media containing 100μg/ml Ampicillin, 15μg/ml Tetracyclin and 0.1% glucose was inoculated with a GlaMor/puc119His6mycXba/TOPP2 clone and the culure was incubated overnight at 30°C, shaking.
2. Reduced (GSH) and oxidised (GSSG) glutathione (SIGMA) was added to a final concentration of 0.2mM each, the pH of the media was adjusted to 6 and the culture was induced with 1mM IPTG at 26°C for 3 hours, shaking.

   As a control an identical culture was set up without the addition of glutathione.
3. The cells were harvested, lysed and the scFv was purified following the expression and purification protocols described previously in this Chapter. GlaMor scFv

fragments were analysed by SDS PAGE electrophoresis and detected by Western Immunoblotting.

Results

As shown in Figure 6.5, no significant effect on the yield of soluble GlaMor scFv fragment was observed, indicating that the expression behaviour was probably not due to the aberrant behaviour of the free cysteine.

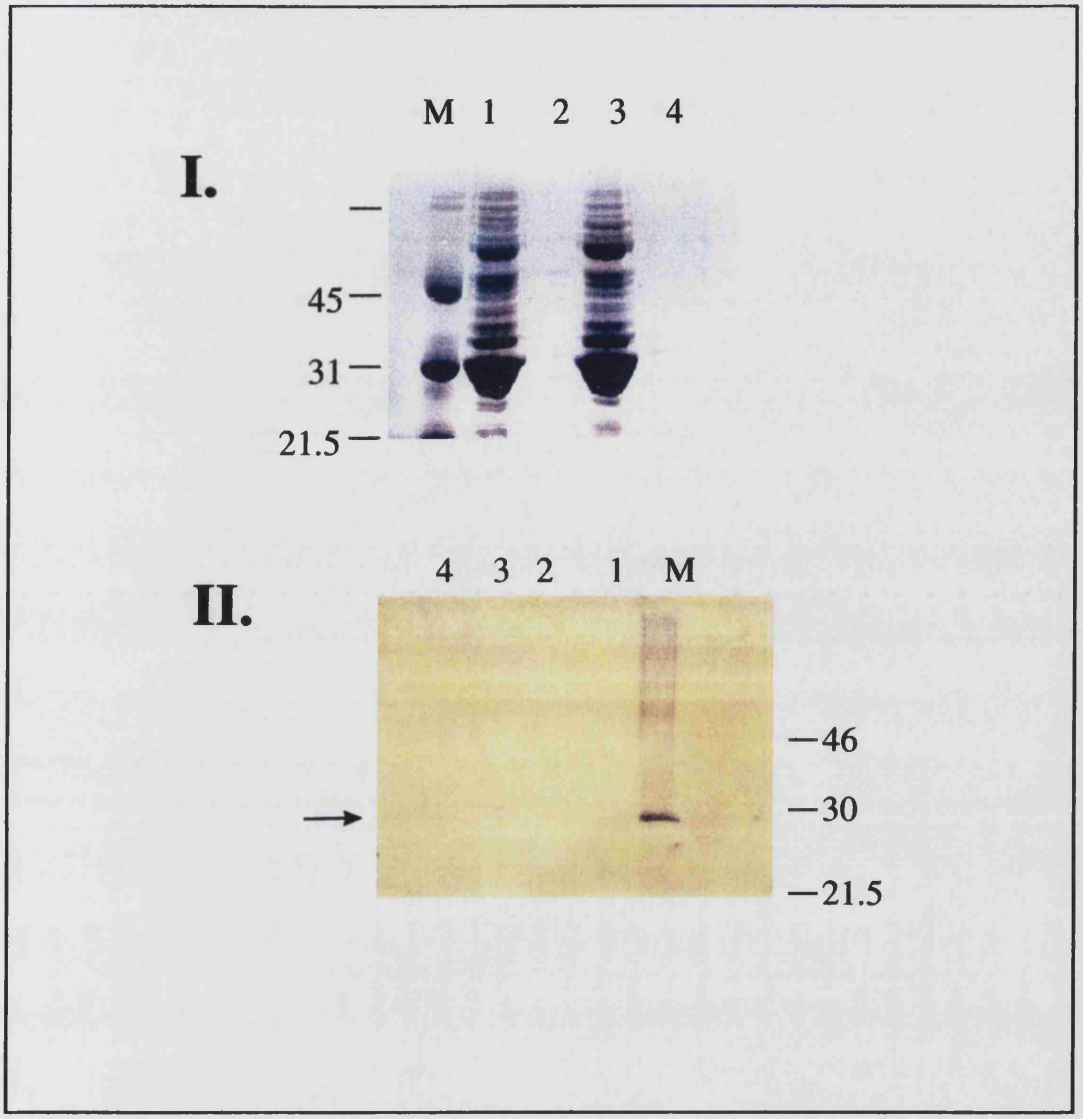


Fig. 6.5: Reducing 12% SDS-PAGE (I.) and Immunoblot analysis (II.) of the periplasmic expression of GlaMor scFv using a redox buffer

0.2mM of each reduced (GSH) and oxidised (GSSG) glutathione was added to GlaMor scFv expression culture just before induction. The periplasmic lysates were subjected to IMAC purification and selected fractions were concentrated and buffer exchanged in PBS. 15µl from each sample was analysed per lane.

I. Proteins were stained with Coomassie Brilliant Blue

1 and 3, lysates before purification with and without glutathione during induction

2 and 4, purified lysates with and without glutathione during induction

II. Western Blot analysis.

Samples 1, 2, 3 and 4 as described above.

The molecular weight markers (M) have protein bands of relative molecular weights of 21.5; 31;and 45 kDa for the SDS-PAGE, and 21.5; 30; and 46kDa for the Western.

6.6 Bacterial expression of Gloop2 and GlaMor scFv.

The studies described in this chapter addressed several problems encountered when periplasmic expression was attempted in *E. coli*, associated mainly with the very low yields of the GlaMor construct. Gloop2 expression levels were also low with 2μg/ml of culture as the maximum concentration obtained. The optimisation trials that have been carried out aimed to maximise the expression yield and the results were consistent for the two constructs with the Gloop2 giving better yields than GlaMor. A lower amount of scFv in the periplasmic or culture supernatant led to less pure material after the IMAC purification. The process of concentrating the fractions obtained was laborious and extended the time of the experiments, which in turn increased the risk of proteolytic degradation.

Attempts to improve the expression yields of GlaMor included the addition of redox buffer in the media during induction as well as sucrose in a separate set of trials. Both methods were based on earlier studies by Proba *et al.*(1997) and Kipriyanov *et al.* (1997) respectively. The addition of a low concentration of redox buffer was tested in order to improve the folding of GlaMor scFv, which might have been affected by the fifth, unpaired cysteine introduced during PCR mutagenesis. The rationale was that normal formation of disulphide bonds between cysteine pairs involves membrane bound proteins that are displaced by the "normal" partners. If the second is missing the unpaired cysteine might form incorrect disulphide bonds with wrong cysteines from other proteins leading to an aggregation-prone product. Glutathione derivatises the unpaired cysteine, does not interfere with affinity purification and can be removed by mild reduction conditions (Samuelsson *et al.*, 1996). However, these studies showed no effect of glutathione on the expression yields of GlaMor scFv.

Sucrose added in the media has been proposed to stabilise the expressed scFv by inhibiting aggregation. The hypothesis is that sucrose increases the osmotic pressure and causes the periplasm to enlarge (Kiefhaber *et al.*, 1991) decreasing the concentration of secreted protein and resulting in less aggregation. The studies reported here showed that although in the Gloop2 system the addition of sucrose increased the yield of the scFv that was directed into the medium after a 24-hour induction by up to 100%, it had no effect on the expression levels of GlaMor.

The large amount of material found of 31kDa in Fig. 6.5 might be attributed to material being trapped in cytoplasmic inclusion bodies. There are methods available that enable the extraction and refolding of inclusion body scFv's. For example, Pantoliano *et al.* (1991). Conformational stability, folding and ligand binding affinity of single chain Fv immunoglobulin fragments expressed in *E. coli.*) describe the expression and renaturation of an anti-fluorescein scFv at levels of 26mg/L. By contrast, this scFv yields almost no protein during periplasmic expression (Jung & Pluckthun, 1997). However, since those early studies it is clear that i) not all scFv's behave in the same manner during expression, either into the cytoplasm or the periplasm, and ii) that refolding from cytoplasmic inclusion bodies has fallen out of favour due to the ease with which secreted (periplasmic) antibody can be produced, and the absense of a single cytoplasmic protocol that will work for all scFv's (J.Huston, pers.comm. and Proba *et al* (1995). Functional single chain antibody fragments from the cytoplasm of *E.coli*: influence of thioredoxin reductase). In fact, the majority of publications in scFv expression since those early experiments in the laboratories of Bird and Huston have almost exclusively centred on periplasmic expression. One of the factors that influences expression in E.coli seems to be the particular residues that occupy certain framework positions that are central to the

refolding of scFv. Extensive studies by the Pluckthun group have established that *E. coli* expression is heavily influenced by sequence (Worn, A & Pluckthun,A, 1998). Mutual stabilisation of VL and VH in single chain antibody fragments, investigated by mutants engineered for stability (Worn & Pluckthun, 1999) Differential equilibrium stability behaviour of scFv fragments: Identification, classification and improvement by protein engineering (Wall,JG & Pluckthun,A, 1999). The heirarchy of mutations influencing the folding of antibody domains in *E.coli.*). The use of alternative expession sysems such as yeast were not considered in-house due to the complexity of evolving a new system in which little experience was held – even expert groups have exerted a great deal of effort to develop such systems (Shusta *et al.*, 1998).

In retrospect, attempting cytoplasmic expression and purification may have been an option that could have been explored. The decision not to do so was a fine balance between knowledge of the periplasmic system already used successfully by other laboratories and members of this laboratory, and the time remaining for the project. In the event, a decision was taken to undertake further *E. coli* studies with the aim of optimising GlaMor expression levels in collaboration with expert colleagues at the Institute of Applied Microbiology, University of Agricultural Sciences, Vienna (Dr. Randolf Kerschbaumer under the supervision of Professor Florian Rueker). These studies are discussed in Chapter 7.

7

Expression Studies on a Critically Unstable scFv[1]

7.1 Introduction

Functional expression yields of antibody fragments in the periplasm of *E. coli* (Pluckthun, 1992), especially Fv or scFv fragments, vary widely over several orders of magnitude even when the cell density is accounted for or when fragments in the identical host-vector system are compared (Carter *et al.*, 1992; Pluckthun *et al.*, 1996). Despite numerous studies (reviewed by Pluckthun *et al.*, 1996), the factors influencing antibody expression are still poorly understood. In this Chapter expression trials on GlaMor scFv are described. Some unusual expression conditions (i.e high glucose media and 16°C induction temperature) that have been previously shown to decrease

significantly plasmid loss and have been used successfully for expression of other scFv's are attempted (Kerschbaumer *et al.*, 1997).

Expression of protein-scFv fusion constructs has been used in cases where stabilisation of the expressed scFv fragment is required and in other studies where the signal peptide has not mediated efficient protein translocation through the inner membrane to the bacterial periplasm. In this study, the GlaMor scFv fragment was fused to the N-terminus of *E. coli* alkaline phosphatase (Kerschbaumer *et al.*,1996).

Sequencing of the pUC119HisMycXba construct identified on GlaMor an additional mutation introduced either during PCR amplification procedures or during expression in the first residue of FW4 of the light chain (position 98) where the originally identified Phenylalanine has been mutated to Leucine. This mutation and its effect on the expression/binding profile of GlaMor scFv is also investigated in this chapter.

In addition, construction of an expression plasmid vector that normally ensures high levels of protein synthesis (Balbas and Bolivar, 1990) will be examined. GlaMor has been recloned into vectors with optimal architectures (optimised promoter, ribosomal binding site (RBS) and Shine-Dalgarno (SD) sequence and initiating AUG codon spacing; Makrides *et al.*, 1996).

Primary sequence effects have also been analysed. It has already been shown that the presence of certain residues at particular positions in the V-region sequence can be a decisive factor in determining the yield of functional protein (Knappik and Pluckthun,

[1]Study carried out by Professor Florian Rueker and Dr. Randolf Kerschebaumer, Institute of Applied Microbiology, Unioverity of Agricultural Sciences, Vienna.

1995 and Nieba *et al.*, 1997). Also, Ulrich *et al.* (1995) found that point mutations in the CDRs can increase the yields in periplasmic antibody fragment expression.

The above factors: expression conditions, fusion partners, expression vectors and scFv primary sequence will be discussed in this chapter as a case study and will provide further confirmation of the unsuitable expression profile of the GlaMor scFv.

7.2 Expression and Purification of GlaMor scFv

Protocol

The pUC119/His/Xba clone of GlaMor (constructed as described in Chapter 6) was used to transform *E. coli* strain TG1. The expression and purification protocol is described in detail in Kerschbaumer *et al.*(1997). According to this protocol the TG1 cell culture is grown in high glucose medium (4g in 1lt). The culture is then centrifuged and the cells are resuspended in fresh medium without glucose. IPTG (1mM final concentration) is added for induction and the culture is shaken at 16^0C for 24hours. Cells are harvested and the periplasmic extract purified using Chelate Sepharose column (Pharmacia, Sweden), charged with Zn^{+2}. Fractions are eluted using a step gradient of imidazole (maximum concentration used was 120mM) and analysed by ELISA. scFv was detected by anti-his or anti-myc 9E10 (Munro and Pelham, 1986) followed by an anti-IgG phosphatase conjugate. Colorimetric development indicated the levels of expression measured against a control Fv.

Observations

Using these conditions, no scFv was detected. Some material was eluted with the imidazole gradient but since the elution behaviour observed with GlaMor was different to that routinely seen with other his-tag containing scFv's (Kerschbaumer *et*

al.,1997) the possibility of an *E. coli* protein binding weakly to the column could not be excluded. ELISA using the anti-his and 9E10 antibodies to detect GlaMor scFv were also negative.

Conclusions

The presence of the *pel*B leader sequence at the N-terminus, as well as the his and myc tags at the C-terminus were verified by sequencing of the GlaMor scFv clone. Protease action might have been responsible for the above observation if. For example, the his and myc tags had been clipped off the recombinant protein. However, the expression was carried out at 16°C and the time of purification minimal, reducing the likelihood of proteolytic action. The control scFv, Gloop2 did not exhibit similar problems.

The "apparent" absence of GlaMor scFv from the periplasmic fraction might also be explained by non-efficient translocation of the molecule through the inner membrane (although the leader sequence is intact) maybe related to folding efficiency and stability.

Further constructions

Selected fusion systems have been shown to have a direct impact on high level production and, in some cases secretion of the target proteins (for a review see Makrides *et al.*, 1996). A simplified method has been designed for fusion of scFv fragments to the N-terminus of *E. coli* alkaline phosphatase (AP) and the resulting immunoconjugates can be produced by expression in *E. coli* and purified in a single step via metal affinity chromatography due to an added his-tag (Kerschbaumer *et al.*, 1996). Application of this method on GlaMor scFv will facilitate its detection through its phosphatase activity, as well as potentially stabilising the scFv (alkaline

phosphatase is very stable and well expressed, Kerschbaumer, personal communication).

7.3 Expression and Purification of Alkaline Phosphatase (AP) Fusion GlaMor scFv

The pDAP2 vector (Fig. 7.1) constructed by Kerschebaumer *et al.* (1996) was designed for simple and rapid construction of scFv -AP fusion proteins. Genes of scFv can be cloned into the polylinker site, thus linking them to the 5' end of the ecphoA1-gene. The *lac*-operator allows induction of expression by IPTG, while the N-terminal *pel*B-leader directs the product to the periplasmic space to allow simple harvesting of concentrated fusion protein. Additionally, a hexa-histidine-tag fused to the C-terminus of AP, facilitates purification by metal affinity chromatography in one step.

Protocol

GlaMor scFv gene was subcloned into the pDAP2 vector using the *Hind*III/*Not*I cloning sites (Fig. 7.1III). The resulting construct maintains the ORF of the pUC119HisXba parent clone, has no *myc* tag and has the old 6x*his* tag replaced by the alkaline phosphatase gene followed by a 6x*his* tag at its C- terminus. The expression and purification was performed as described by Kerschbaumer *et al.* (1996). All cytoplasmic, periplasmic and culture supernatant fractions were purified by metal affinity chromatography, eluted with imidazole (1mM final concentration) and tested for AP activity with p-nitrophenyl as substrate, the active fractions were pooled and analysed on SDS-PAGE followed by silver staining (Fig. 7.2).

Observations

Using AP activity to trace AP-GlaMor scFv showed that it was located in the cytoplasmic fraction, despite the *pel*B leader sequence being present. The purified AP-GlaMor scFv fragment from the cytoplasmic fraction showed as a single band approximately 8-10kDa smaller than the expected size of 78kDa (Fig. 7.2). For comparison, expression results using other scFv's are also shown.

Conclusions

Various possible explanations of the above result can be proposed. For example, a premature stop codon may have been introduced during the vector construction. However, the fact that the protein can purify using the his-tag as well as extensive sequencing checks argue against this.

The most probable explanation is that the cytoplasmic location of the fusion protein exposed it to proteolytic degradation - the cytoplasm of *E.coli* contains a greater number of proteases than the periplasm. (Swamy and Goldberg, 1981 and 1982). Strategies for minimising proteolysis of recombinant proteins in *E. coli* include protein targeting to the periplasm or the culture medium, use of protease deficient host strains, growth of host cells at low temperature and construction of N-and/or C-terminal fusion proteins (see Makrides *et al.*, 1996 for a review). However, Wall and Pluckthun (1995) have suggested that when using protease-deficient strains,
low efficiency expression may result due to the accumulation of abnormal proteins and the resulting toxicity.

Again, the configuration of the procaryotic expression vector has been suggested as a key factor that influences the levels of protein synthesis (Balbas and Bolivar, 1990). The essential architecture of an *E. coli* expression vector needs to fulfil some

minimum requirements: the promoter is positioned approximately 10-100bp upstream of the ribosome-binding site (RBS) and it is under the control of a regulatory gene. The RBS spans a region of approximately 54 nucleotides and includes part of the promoter, the shine-Dalgarno (SD) sequence, part of the coding sequence and the spacings between them. The SD sequence interacts with the 16S rRNA during translation initiation. The spacing between the SD site and the initiating AUG codon ranges from 5 to 13 nucleotides and influences the efficiency of translational initiation (Gold, 1988). The distance between the LacZ-RBS and the ATG of the *pel*B leader sequence in GlaMor pUC119HisXba and in initial pDAP2 constructs is somewhat longer than 13 nucleotides.However, when GlaMor scFv was recloned into the pDAP2 vector at *Sfi*I/*Not*I maintaining the ORF of the pDAP2 vector, no improvement was seen.

Future Constructions

One possible cause of the low expression behaviour, not yet examined, is the presence of the non-designed mutation, Phenylalanine to Leucine, at the start of FW4 of the light chain. This position is highly conserved among all FW4 light chain antibody sequences and is not present in the Gloop2 parent sequence whose expression is more normal. It has been suggested (Knappik and Pluckthun, 1995) that the primary sequence of a particular antibody variable region emerges as the most decisive factor determining the yield of functional protein and that very minor changes can have a dramatic effect on the *in vitro* aggregation properties of these molecules.

The back mutation of Leucine to Phenylalanine and the use of the pDAP2 vector with its optimal configuration are discussed in the next section.

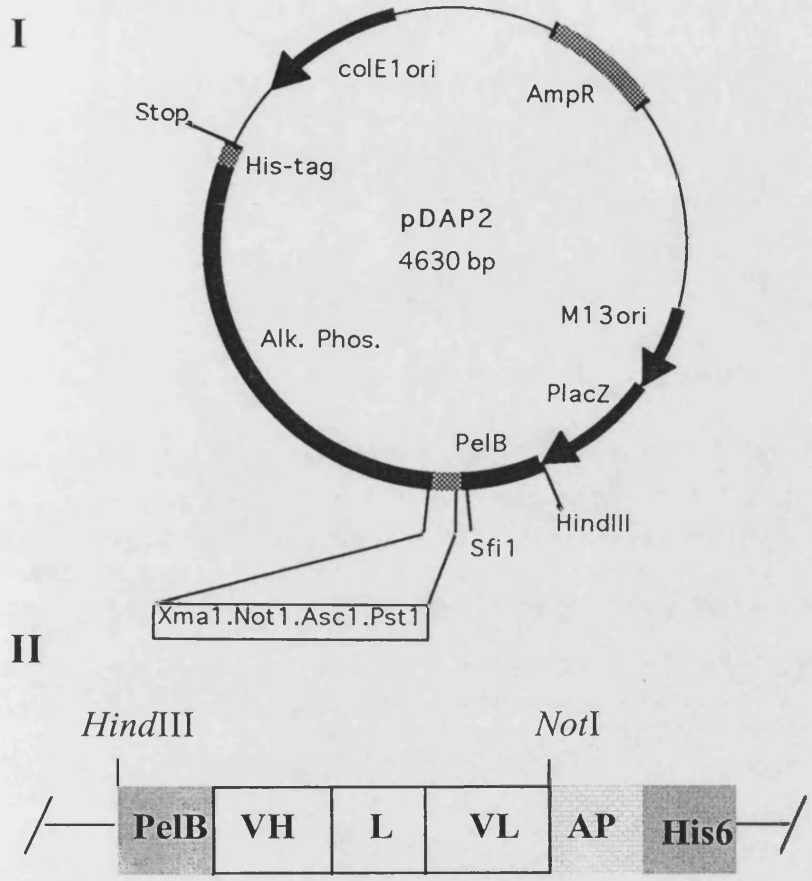


Fig. 7.1: Cloning of GlaMor scFv in pDAP2 vector

I. Map of the pDAP2 vector (Kerschabumer *et al.*, 1996)

II. GlaMor scFv cloning cassette in the pDAP2 vector

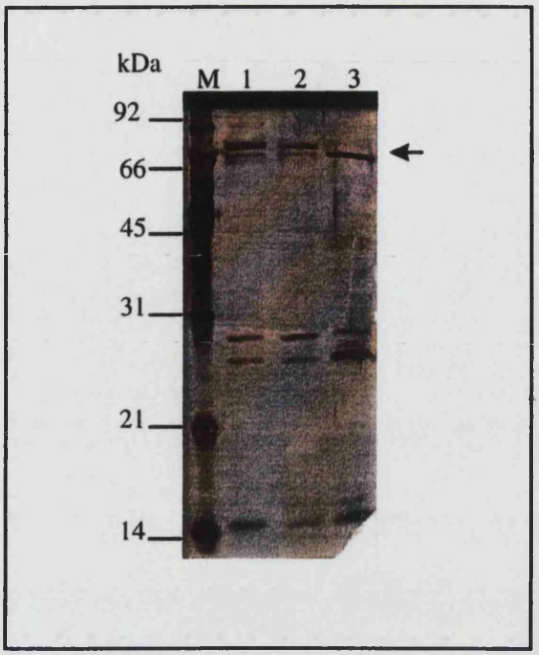


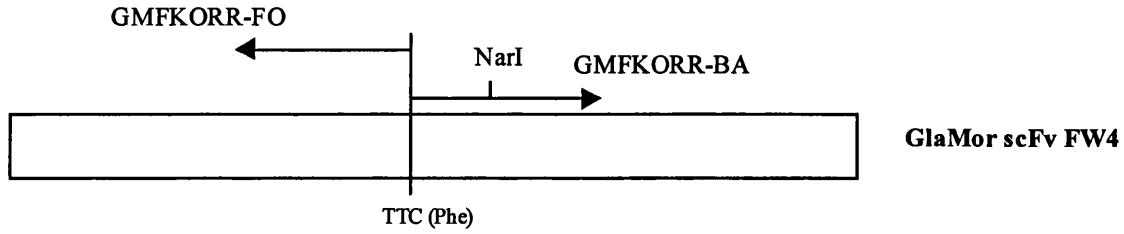
Fig. 7.2: SDS PAGE of AP-GlaMor scFv purified periplasmic extract (3).

Silver staining gel showing the purified periplasmic fraction in lane 3 and other control expressed scFv's in lanes 1 and 2. Proteins of known molecular weight are shown in lane M

7.4 Back Mutagenesis of Leucine to Phenylalanine using PCR

Protocol

Two primers and inverse PCR were used for making the leucine to phenylalanine mutation of the GlaMor FW4 light chain. The sequences of the two primers GMFKORR-BA and GMFKORR-FO as well as the region on the scFv that these prime are shown below:



GMFKORR-FO: 5'CGTGTCCGGA AAACTAAGAT CTTGG

GMFKORR-BA: 5'**TTC**GGCGCCG GGACCAAGCT CGAGCTGAAA C

The GMFKORR-BA primer contains the correct **TTC** phenylalanine codon, the *NarI* site and primes clockwise while the GMFKORR-FO primer primes anticlockwise. The *NarI* site (shown underlined) was introduced as a silent mutation to facilitate screening. Inverse PCR using this set of primers results in a linear PCR product which can be closed by ligation. The respective ends of the primers when ligated together as blunt ends give the correct reading frame. DNA from the resulting clones was screened for digestion with *NarI* and positive clones were thereafter sequenced. DNA from six clones was prepared and sequenced. Each of the six clones was found to contain either the old mutation or different mutations (insertions, deletions or duplications). All six clones and the mutations identified in each are listed below.

Observations

Clone1

GMSEQ3: has retained the old mutation

```
                                                    ▼
GLAMOR    - CAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGTTCGGTGCT -750
            ||||||||||||||||||||||||||||||||||||||||| |||| ||
GMSEQ3_IC - CAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGCTCGGCGCC -227

GLAMOR    - GGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACTCATCTC -800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
GMSEQ3_IC - GGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACTCATCTC -277
```

Clone2:

GMSEQ5: contains the correct mutation but lacks one base upstream

```
                                                  ▼
GLAMOR    - TTTGCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGTTCGG -746
            |||||||||||||||||||||||||||||||||||| ||||||||||||||
GMSCFV5_IC- TTTGCAGACTATTATTGTTGCCAAGATCTTAGTTT-CCGGACACGTTCGG -542

GLAMOR    - TGCTGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACTCA -796
              || ||||||||||||||||||||||||||||||||||||||||||||||
GMSCFV5_IC- CGCCGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACTCA -592
```

Clone 3:

GMSEQ8: retains the old mutation and contains a single base insertion upstream

```
                                                  ▼
GLAMOR    - TTTTGCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACAC-GTTC -744
            |||||||||||||||||||||||||||||||||||||||||||| | ||
GMSEQ8    - TTTTGCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACTGCTC -365

GLAMOR    - GGTGCTGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACT -794
            || || ||||||||||||||||||||||||||||||||||||||||||||
GMSEQ8    - GGCGCCGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACT -415
```

Clone 4:

GMSEQ9: retains the old mutation and an additional point mutation downstream

```
GLAMOR    - TTGCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGTTCGGT -747
            |||||||||||||||||||||||||||||||||||||||||||| ||||
GMSEQ9_IC - TTGCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGCTCGGC -402
                                                  ▼
GLAMOR    - GCTGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACTCAT -797
            || |||||||||||||||||| |||||||||||||||||||||||||||
GMSEQ9_IC - GCCGGGACCAAGCTCGAGCTGAACCGGGCGGCCGCAGAACAAAAACTCAT -452
```

Clone 5:

GMSEQ10:conatins the correct mutation, but has duplicated bases downstream (126 bases)of GlaMor (see complete sequencing result for GMSEQ10 below)

```
GLAMOR     - GCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGTTCGGTGC -749
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||
GMSEQ10_IC- GCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGTTCGGCGC -409

GLAMOR     - TGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACTCATCT -799
              |||||||||||||||||||||||||||||||||||||||||||||||||
GMSEQ10_IC- CGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAACTCATCT -459
```

Dublicated Bases (shown in italics and underlined)

CCAGTTNCAT CTCNTATGGC CTCTNTGGAA AANGTCAGTC TCACTGTGGG CAGTCAAGAA
ATTAGTGGTT CTTAAGCTGG CTTGGCAGAA CCAGATGAAC TATTAAACGC CTATCTACGC
NGCATCCACT TTAGATTCTG GTGTCCCAAA AAGGTTCAGT GGCAGAAGGT CTGGGTCAGA
TTATTACTCA *CCATCAGCAA GCCTTGAGTC TGAAGATTTT GCAGACTATT ATTGTTGCCA* ◀
AGATCTTAGT TTTCCGGACA CGTTCGGCGC CGGGACCAAG CTCGAGCTGA AACGGCGGCC
*GCAGAACAAA AACTCA*GATT ATTCACTCAC CATCAGCAAG CCTTGAGTCT GAAGATTTTG ◀
CAGACTATTA TTGTTGCCAA GATCTTAGTT TTCCGGACAC GTTCGGCGCC GGGACCAAGC
TCGAGCTGAA ACGGGCGGCC GCAGAACAAA AACTCATCTC AGAAGAGGAT CTGAATGGGG
CCGCACATCA CCATCATCAC CATTAATAAG AATCCANTGG CCGTCNTTAT ACAACGTCGT
GGCTNNAAAA A

Clone6:

GMSEQ11: contains additional mutatons around the *NarI* site

```
GLAMOR     - CTGGGTCAGATTATTCACTCACCATCAGCAGCCTTGAGTCTGAAGATTTT -699
             ||||||||||||||||||||||||||||||||||||||||||||||||||
GMSEQ11_IC- CTGGGTCAGATTATTCACTCACCATCAGCAGCCTTGAGTCTGAAGATTTT -353
                                                    ▼  ▼  ▼
GLAMOR     - GCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGTTC-GGTG -748
             |||||||||||||||||||||||||||||||||||||||||||| |  || |
GMSEQ11_IC- GCAGACTATTATTGTTGCCAAGATCTTAGTTTTCCGGACACGCTATGGCG -403
             ▼
GLAMOR     - CTGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAA -797
             | ||||||||||||||||||||||||||||||||||||||||||
GMSEQ11_IC- CCGGGACCAAGCTCGAGCTGAAACGGGCGGCCGCAGAACAAAAA -453
```

Conclusions

Inverse PCR worked to repair Leucine back to Phenylalanine. However, it introduced other mutations into the GlaMor scFv gene. Although, the PCR process itself might be causing this problem, some instability deriving from the GlaMor scFv gene cannot be excluded. This is particularly strengthened by the fact the Pfu polymerase (Stratagene) used for these experiments has a much lower error rate and is unlikely to have been responsible for the observed mutations.

The above GlaMor sequences were screened (50bp-1bp) either side of the problematic region against the EMBL database, to identify possible matches with other unstable sequences of other 'difficult' Fv's. The search revealed antibody-related sequences none of which have been catalogued as giving rise to instability or problematic expression.

7.5 Non-PCR Back Mutagenesis of Leucine to Phenylalanine.

Protocol

GlaMor/pUC119his6mycXba was double digested with *BglII* and *Not*I and the large piece of DNA was gel purified. The two primers:

- GM-KORR-CO (56mer): *gat ct*t agt ttt ccg gac acg **ttc** ggc gcc ggg acc aag ctc gag ctg aaa cgg gc and
- GM-KORR-IC (56mer): *ggc cgc* ccg ttt cag ctc gag ctt ggt ccc ggc gcc gaa cgt gtc cgg aaa act aa

when annealed yield *Bgl*II and *Not*I overhangs (in italics) and were ligated with the cut plasmid (Fig.7.3) This step replaced the piece containing the fault with a synthetic correct sequence (in bold). Positive clones were selected and characterised by restriction digest, using *Bgl*II and *Not*I to cut out the insert and a newly created *Nar*I site (unerlined) contained on the insert itself (silent mutation to facilitate screening). Positive clones were verified by sequencing the whole GlaMor sequence.

The GlaMor scFv gene was subcloned in the pDAP2 (Fig. 7.1). Briefly, pDAP2 vector was digested with *Pst*I and *Not*. The GlaMor gene was prepared by digesting the gene coding for the repaired GlaMor with the same enzymes and purifying using preparative gel electrophoresis. Then, the gene was inserted into pDAP2 and the construct was verified by restriction analysis. Positive constructs were used for expression as described in paragraph 7.3 It should be noted, that bacteria harboring GlaMor constructs were grown at less than 25°C, for both DNA preparation or protein expression to minimise the instability risk.

The expression and purification were performed as follows:

E. coli strain TG1 harboring the GlaMor/pDAP2 construct were grown up in M9ZB containing 100µg/ml ampicillin and 2%(w/v) glucose to an OD_{600} of 0,8-0,9. Cultures were cooled down to 16°C, IPTG was added to 1mM and cultures were shaken for 16-18 hours at 16°C. Cells were harvested by centrifugation (11,000 x *g*, 4°C, 10min), resuspended in 0.2M borate pH 8.0 containing 1mMNaCl and 1mM EDTA and incubated on ice for 30min.

After centrifugation (13000 x *g*, 4°C, 20min) the supernatant was recovered and the pH was adjusted to 7.35 by adding 1M KH_2PO4. Imidazole was added to 1mM final concentration. The phosphatase activity was determined in 10µl sample by adding 100 µl substrate buffer (1mgPNPP/ml, 50mM Tris pH10, 150mM NaCl, 2mM $MgCl_2$), incubation for 45min at room temperature and measuring the OD at 405nm. Purification was performed by using *HiTrap* Chelating Sepharose:column (Pharmacia, Sweden). The column was loaded with 0.1M $ZnCl_2$ and equilibrated with PBS/1M NaCl. The periplasmic fraction was loaded on the column, then washed with equilibration buffer and eluted in 1ml fractions by increasing imidazole concentrations of 10mM up to 120mM.

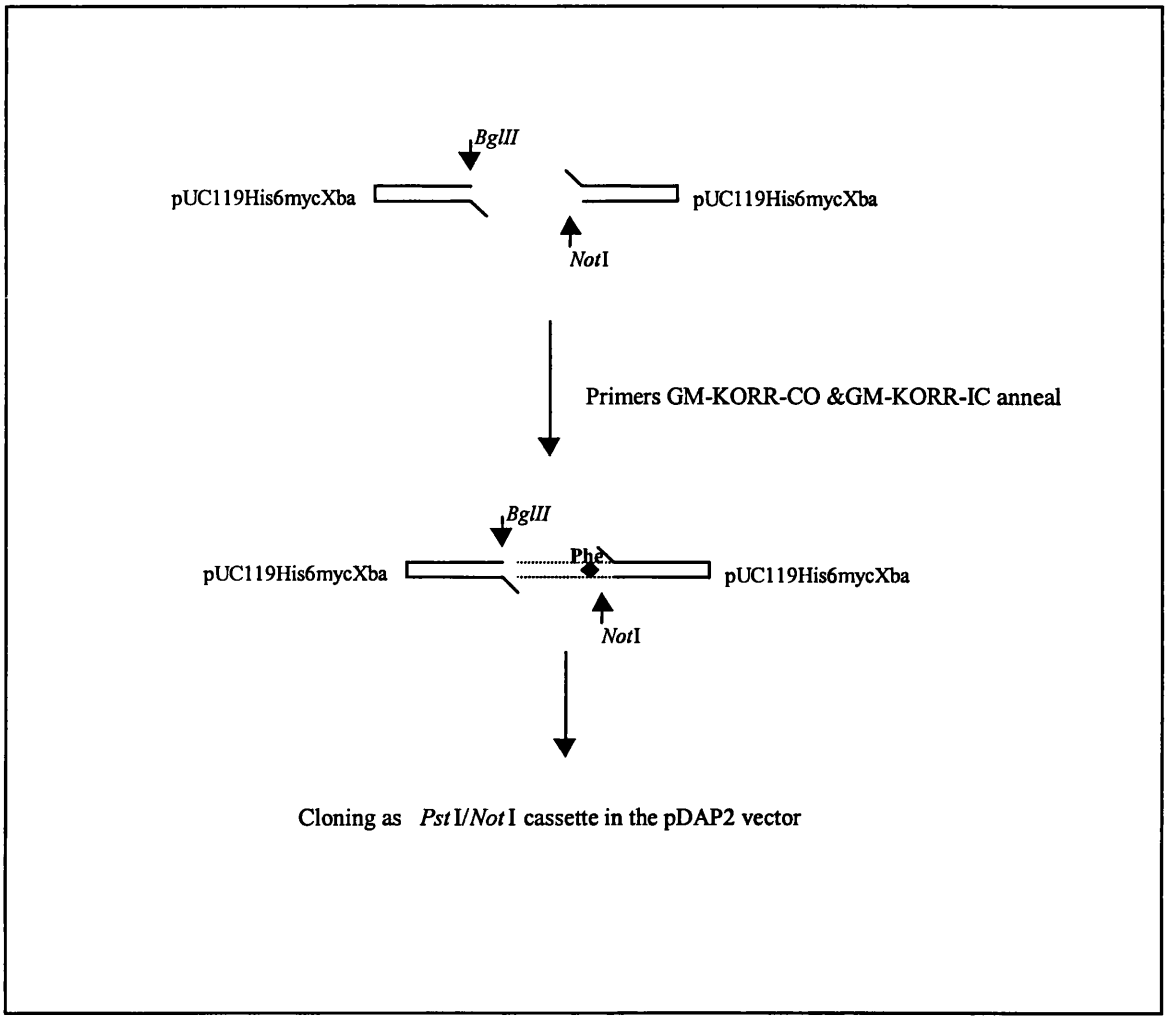


Fig.7.3: Non-PCR mutagenesis of Leucine back to Phenylalanine.

Observations

The OD_{405} measurements are shown at the table below.

Before IMAC Purification		After IMAC Purification	
M9ZB Blank	0.032	Periplasmic Extract	0.088
Buffer Blank	0.024	Flow Through	0.420
Supernatant	0.117	Wash (without Imidazole)	0.038
Periplasmic Extract	0.194	10mM Imidazole	0.039
		20mM Imidazole	0.080
		20mM Imidazole	0.204
		20mM Imidazole	0.226*
		40mM Imidazole	0.245*
		60mM Imidazole	0.657*
		80mM Imidazole	0.844*
		120mM Imidazole	0.044

The four fractions marked with asterisks were pooled and frozen in liquid N2 at –80°C and analysed by silver staining (Fig. 7.4A). The protein corresponding to MW 75kD had no phosphatase activity, while phosphatase activity was present in a band at 50kDa, possibly corresponding to alkaline phosphatase alone. After Western Blotting with anti-his tag antiboly (Fig.7.4C) both lanes gave a signal at the same MW of 50kDa which was too low to be intact phosphatase-fusion protein, suggesting degradation.

Conclusions

The non-PCR method used for correcting leucine back to phenylalanine succeeded to produce clones with the correct sequence avoiding other unwanted mutations. GlaMor scFv clones with the correct sequence were tried for expression but failed to provide scFv material. The hypothesis that the mutation of Phenylalanine to Leucine might have caused the problems with GlaMor scFv expression is therefore considered unlikely.

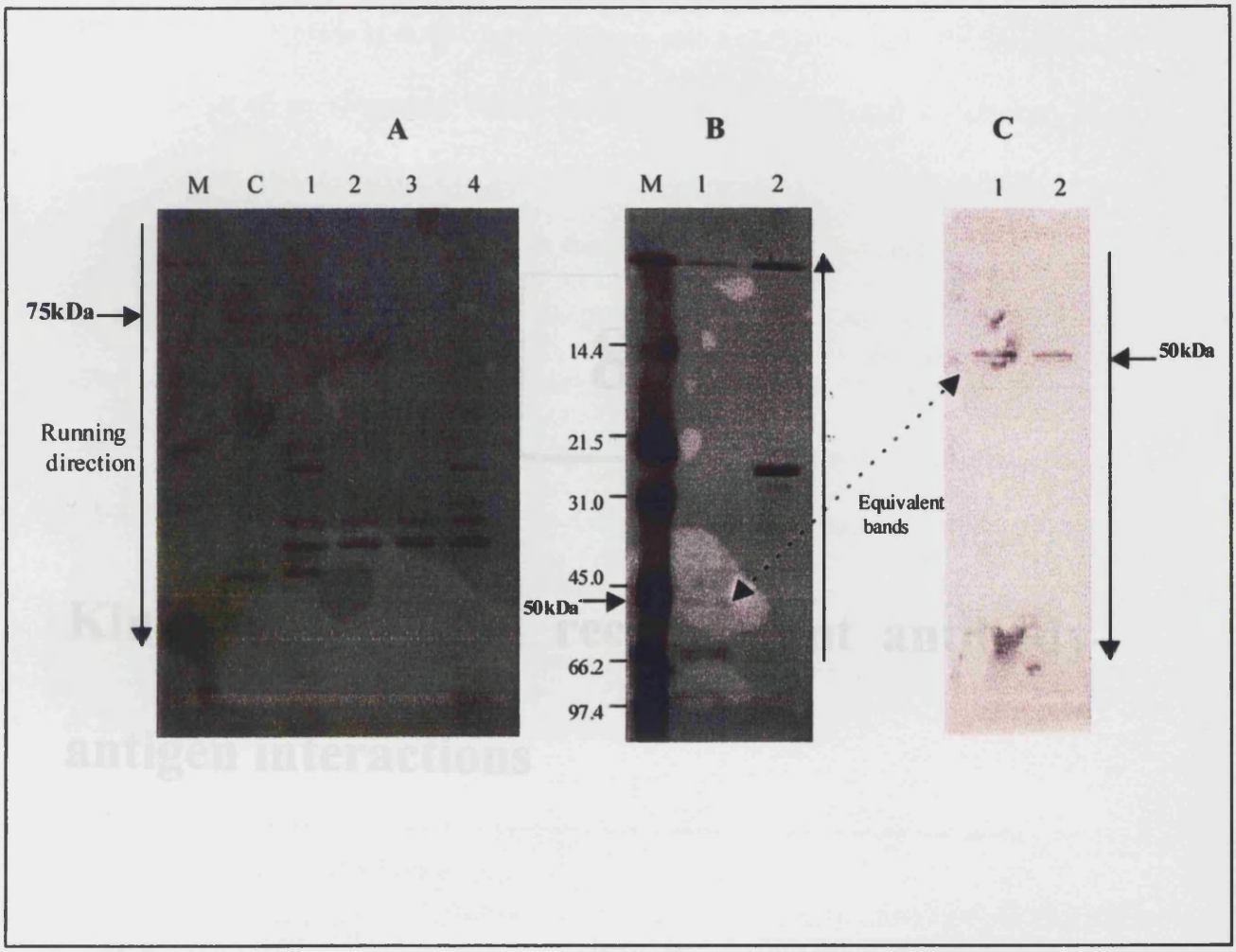


Fig. 7.4: Expression and purification of the repaired GlaMor scFv

Silver staining gel (A) with different fractions (lanes 1-4) after IMAC purification. Proteins of the correct molecular weight are shown at 75kDa. Two of these fractions (A2 and A3), which showed bands of the correct size and had phosphatase activity were pooled with their left and right neighbor (A1and A4) and are shown on gel B (silver staining gel) and gel C (anti-his tag western blot).

8

Kinetic analysis of recombinant antibody-antigen interactions

8.1 Introduction

The relationship between domain structures of recombinant monoclonal antibody fragments and their reaction kinetics was studied for the first time using a biosensor based on surface plasmon resonance (SPR) by Borrebaeck *et al.* in 1992.

Evaluating newly engineered antibodies requires the measurement of their affinity constants for a given antigen. The biosensor based technology permits real-time mass measurements using either surface plasmon resonance (SPR, BIAcore) or a resonant mirror (Iasys, Fisons). In this chapter the preliminary results described were obtained from the SPR BIAcore™ system (BIAcore).

The BIAcore™ system is built of a processing and a computer unit. The processing unit consists of an integrated fluidic cartridge, a sensorchip and a detector. The interaction between biomolecules takes place on the sensor chip, which consists of an optically flat glass slide covered with a thin layer of gold on one side. A hydrophilic dextran matrix is covalently bound to the gold film. (Fig. 8.1). Antibody or antigen is coupled to the matrix and allowed to react with its counterpart. The amount of protein bound to the gold surface is analysed by measuring the surface plasmon resonance (SPR). A change of 0.1 degree corresponds to 1ng bound protein/mm^2 or 1000RU (RU= resonance unit which is an instrument specific unit). The SPR results are illustrated as a sensorgram, where resonance units (RU) are plotted against time (Karlsson *et al.*, 1991). For further details see Appendix 2.

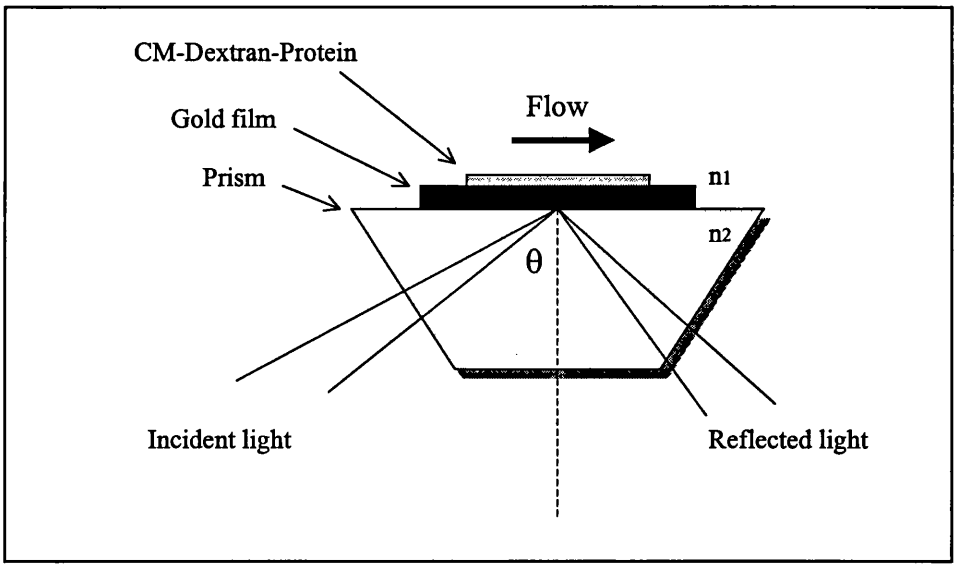


Fig. 8.1: Schematic of biosensor optics
In SPR when a resonant coupling between the incident light energy and surface plasmons in the conducting film occurs at a specific angle of incident light, absorbing the light energy and causing a characteristic drop in the reflected light at that angle.

8.2 Materials and Methods

BIAcore Sensor chip CM5 and the amine coupling kit containing NHS, EDC and a solution of 1M ethanolamine hydrochloride adjusted to pH 8.5 with sodium hydroxide were obtained from BIAcore. The studies described on this chapter were carried out on a BIAcore 2000 at BIAcore (Stevenage, U.K).

8.2.1 Immobilisation of 9E10 mAb on CM5 sensor chip

The general procedure for the conditioning of the sensor chip CM5 surface using HBS buffer was carried out before the activation of the carboxyl groups as described by the manufacturer (BIAcore). After the sensor chip was conditioned, the automated immobilisation cycle was performed at a flow rate of 20μl/min. 70μl each of 0.4M EDC and 0.1M NHS were mixed together and injected over the surface. 9E10 mAb (50μg/ml) was immobilized on all four flowcells of the Biacore 2000 system, in 10mM acetate buffer at pH 5.0 at a flow rate of 10μl/ml. Standard HBS buffer (10mM Hepes pH 7.4, 150mM NaCl, 3.4 mM EDTA, 0.005% Surfactant P20) was used as the running buffer. The procedure was repeated one more time for all flow cells and a third time for flow cell four so as to obtain similar immobilisation levels of 9E10 (in RU) in the four flow cells. Any remaining reactive esters were inactivated by injection of 1M ethanolamine hydrochloride at pH 8.5.

7,600 to 9,500 RU of 9E10 was immobilised on the carboxymethylated dextran surface of the CM5 (BIAcore) on the four flowcells (Fig. 8.2).

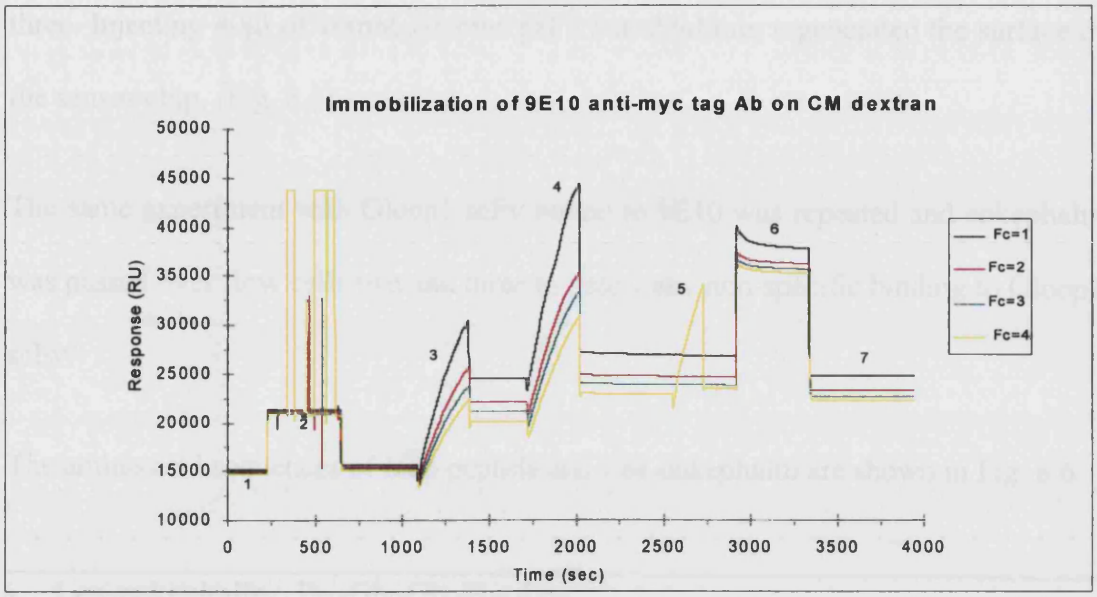


Fig. 8.2: Immobilisation of 9E10 anti-myc tag antibody on the CM sensor chip.
1. baseline of unmodified CM5 sensor chip surface with continuous flow of HBS buffer
2. injection of 140ul of NHS/EDC and activation
3. 1st injection of 50ul 9E10 (50ul/ml) in acetate buffer pH 5.0 over flowcells 1-4
4. 2nd injection (as in 3) over flowcells 1-4
5. 3rd injection (as in 3) over flowcell 4
6. ethanolamine capping of the unreacted esthers
7. immobilised level of 9E10 on all four flowcells (9,500 RU on FC1, 8,300 RU on FC2, 7,800 RU on FC3 and 7,600 RU on FC4.)

8.2.2 Interaction analysis

Gloop2 and GlaMor scFv binding to the 9E10.

Flow cell two (FC2) was kept as a negative control or blank for each experiment. Gloop2 scFv (2μg/ml) was injected at 5μl/min in HBS buffer over flowcell three.1,400RU of Gloop2 was captured on 9E10. (Fig. 8.3). Similarly GlaMor scFv was injected at 5μl/min over flow cell three. Very little of this ligand was available for capture, typically less than 50RU. (Fig. 8.4).

Peptides

40µl of Loop peptide (9µM) in HBS was injected at 20µl/min over flowcells two and three. Injecting 40µl of 10mM Glycine pH 2.2 at 20µl/min regenerated the surface of the sensor chip. (Fig. 8.5).

The same experiment with Gloop2 scFv bound to 9E10 was repeated and enkephalin was passed over flow cells two and three to detect any non-specific binding to Gloop2 scFv.

The amino acid sequences of loop peptide and Leu-enkephalin are shown in Fig. 8.6

- **Leu-enkephalin:** Tyr-Gly-Gly-Phe-Leu
- **Loop peptide:** Gln-Ile-Asn-Ser-Arg-Trp-Trp-Cys-Asn-Asp-Gly-Arg-Thr-Pro-Gly-Ser-Arg-Asn-Leu-Ala-Asn-Ile-Pro-Cys-Ser-Ala-Leu-leu

Fig. 8.6: Amino acid sequences of Loop and enkephalin peptides.

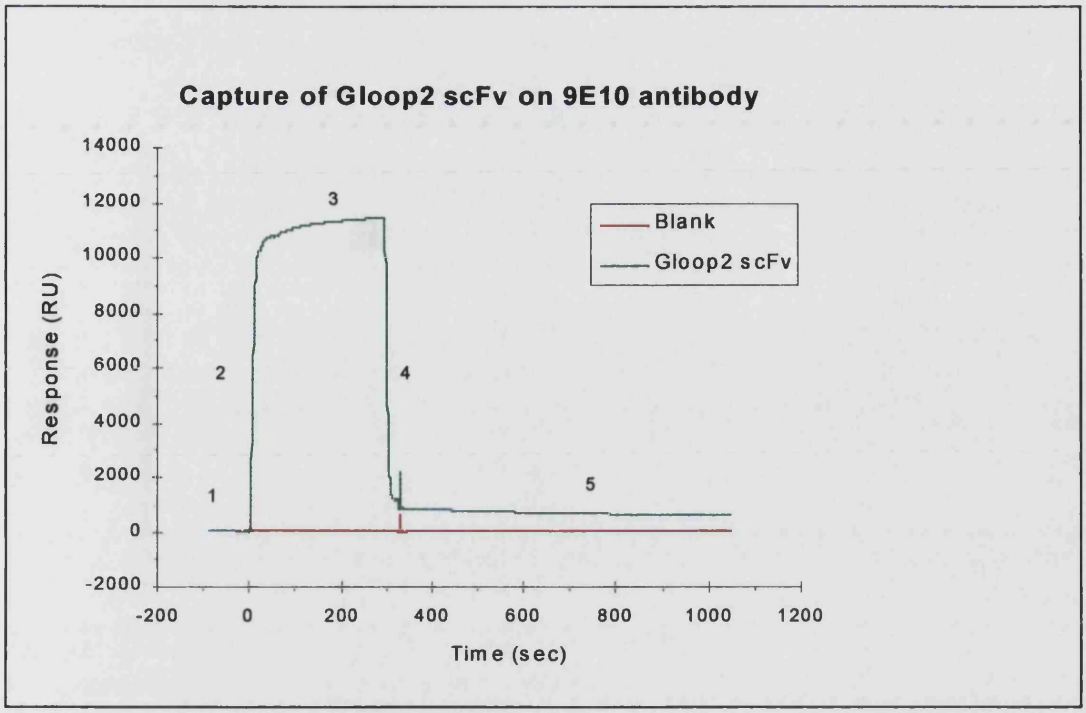


Fig. 8.3: Capture of Gloop2 scFv on 9E10 antibody

1. Shared baseline showing the level of immobilised 9E10 on flowcells 2 and 3 (red and green respectively).
2. Capturing of Gloop2 scFv on flowcell 3 (green)
3. Equilibrium state of Gloop2 on 9E10
4. Wash with HBS buffer
5. Captured levels of Gloop2 scFv on 9E10; approximately 1,400RU of Gloop2 was captured.

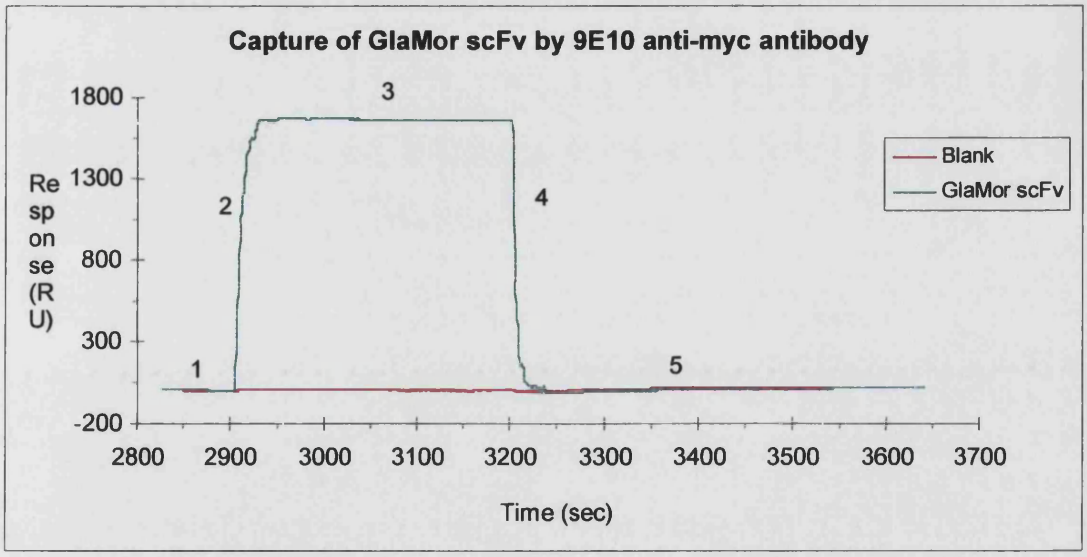


Fig. 8.4: Capture of GlaMor scFv on 9E10 antibody

1. Shared baseline showing the level of immobilised 9E10 on flowcells 2 and 3 (red and green respectively).
2. Capturing of GlaMor scFv on flowcell 3 (green)
3. Equilibrium state of GlaMor on 9E10
4. Wash with HBS buffer
5. Captured levels of GlaMor scFv on 9E10

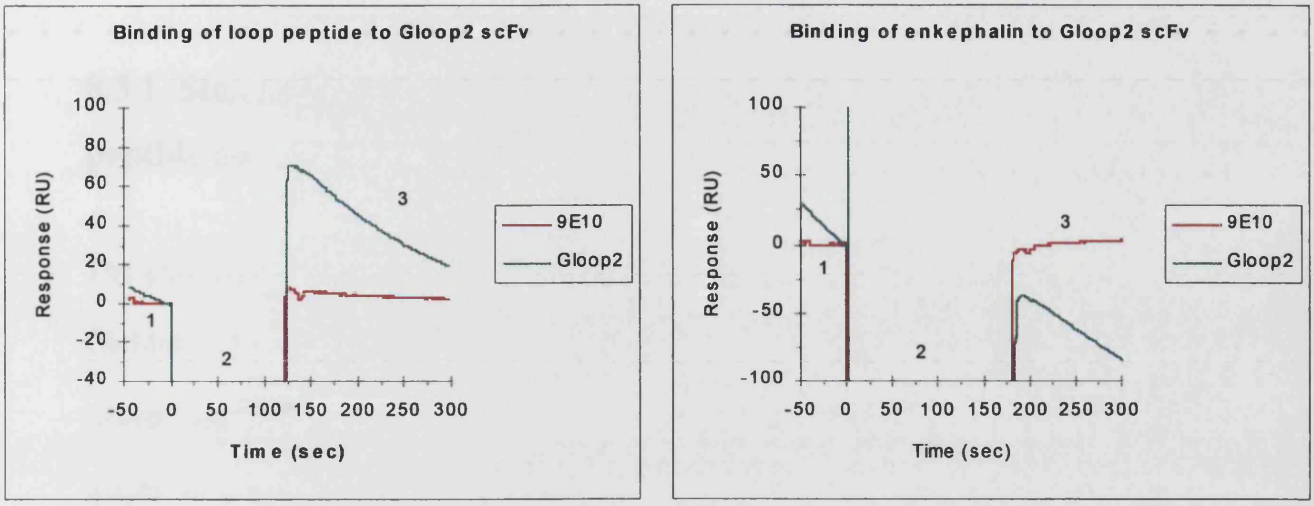


Fig. 8.5: Binding of peptide antigens on Gloop2 scFv

1. Baseline of immobilised 9E10 (blank surface, FC2, red) and captured Gloop2 scFv on 9E10 (FC3, green)
2. Decrease in the SPR signal due to the change of bulk refractive index
3. Wash under continuous flow of HBS of the bound loop peptide. Approximately 50 RU of peptide have been bound on Gloop2 scFv (left sensorgram) while non-detectable amount of enkephalin binds to Gloop2 scFv (right sensorgram).

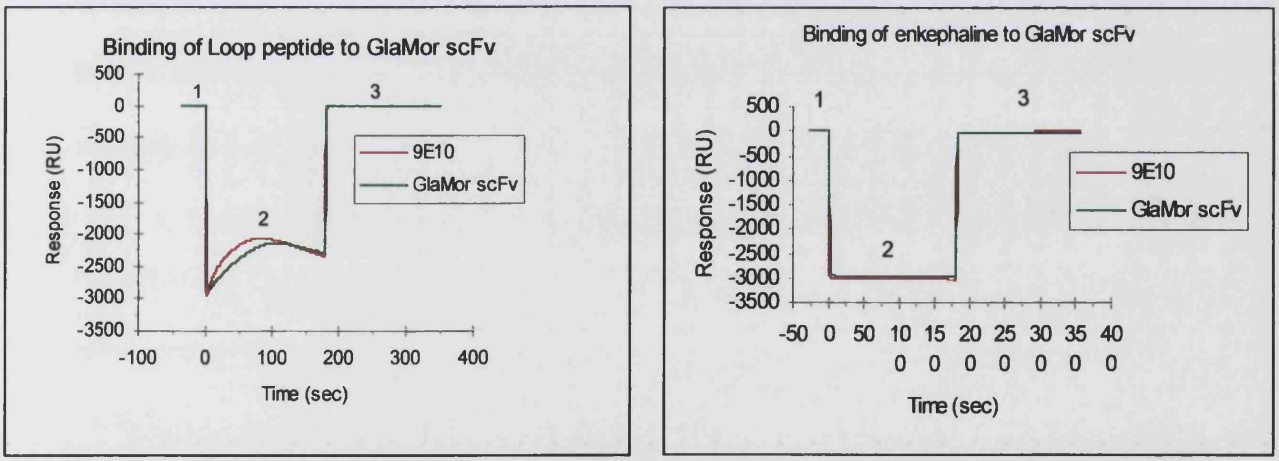


Fig. 8.7: Binding of peptide antigens on GlaMor scFv

4. Baseline of immobilised 9E10 (blank surface, FC2, red) and captured GlaMor scFv on 9E10 (FC3, green)
5. Decrease in the SPR signal due to the change of bulk refractive index
6. Wash under continuous flow of HBS of the bound loop peptide. Non-detectable amounts of either loop peptide (left sensorgram) or enkephalin (right sensorgram) bound on GlaMor scFv.

8.3 Qualitative measurement of binding kinetics

8.3.1 Stoichiometric calculation of maximal binding levels of loop peptide on Gloop2 and GlaMor scFv

The molecular weight of the 9E10 antibody is 150kDa and that of the Goop2 and GlaMor scFv is approximately 32kDa. Hence, the maximal binding capacity of Gloop2 and GlaMor scFv on a sensor surface area with 7,800 RU of captured 9E10 would be 3,120 RU, taking into account 1:2 stoichiometry of binding. For Gloop2 scFv, approximately 1,400 RU was captured using 2μg/ml, which was the highest concentration available. Since the Loop peptide is 1.1kDa the maximal binding capacity of this surface assuming 1:1 stoichiometry would have been 48 RU. For the Loop peptide 50 RU was captured using the 9μM concentration, which was estimated as saturating. Had the Gloop2 scFv shown any specificity for the enkephalin peptide, which has a molecular weight of 500Da, the Rmax (maximum binding capacity) would have been about 20 RU.

A typical binding curve on Biacore may run from 20-2000RUs. For GlaMor scFv in order to detect a 20RU signal of enkephalin bound on GlaMor scFv a minimum of 1400RU of captured scFv is needed. This means that the concentration of expressed and purified GlaMor scFv required should be 2μg/ml. Given the very low amounts of GlaMor scFv expressed and captured on 9E10 no binding of enkephalin on GlaMor scFv could be detected.

8.3.2 Specificity of interaction

The loop peptide bound to Gloop2 scFv with 50 RU when injected at 9μM while no binding of the enkephalin was observed at 9μM injection of the peptide over the Gloop2 scFv. The surface could be effectively regenerated with using 10mM glycine at pH 2.2 for repeated studies on the same sensor surface.

8.3.3 Problems with quantitative analysis

The quantitative analysis that would determine the affinity values of the above interactions was not performed. Low levels of GlaMor scFv expression did not produce enough material to perform complete quantitative analysis. Provided that enough GlaMor scFv was available and specific binding was detected on enkephalin against the loop peptide, assays would be repeated with 5-6 different concentrations of scFv's in order to determine the Kon/Koff and affinity values.

Due to the low molecular weight of the loop and enkephalin peptides all interactions can only be monitored on Biacore 2000 instrument (available in BIAcore, Stevenage) which sensitivity detects low molecular weight analytes.

8.4 Discussion

The Biacore technology offers a non-label tool for studying the antibody-antigen interactions in real time. The experiments described in this chapter aimed to investigate the specificity of Gloop2 scFv and GlaMor scFv's against the loop and enkephalin peptides. The design of the assays took into account the low molecular weight of the peptide antigens and the very low expression yields of the GlaMor scFv as described in Chapter 6. Low molecular weight antigens used as analytes need to be detected by a high sensitivity instrument such as the BIAcore 2000. Alternatively, BSA conjugated forms of the antigens could increase the mass of the analytes but also might interfere with the binding epitope of such small molecules.

The 9E10 anti-myc antibody used for the immobilisation on the CM5 dextran had been used successfully for the detection of periplasmic Gloop2 and GlaMor scFv's in Western Immunoblotting (seeChapter 6). Gloop2 and GlaMor scFv were expressed and purified as described in Chapter 6 and buffer-exchanged in PBS before the assays. Between 1,000 and 1,400 RU of Gloop2 scFv was captured on immobilised 9E10 antibody over several trials that were carried out. This level of immobilised Gloop2 scFv on 9E10 was sufficient to proceed with the binding of the loop and enkephalin peptide antigens. The assays carried out showed specific binding of the Loop peptide on Gloop2. The enkephalin peptide did not show specific binding when used at the highest (9μM) concentration. The GlaMor scFv expressed at very low levels and, when immobilised on 9E10, calculations indicated that insufficient peptide would be captured to allow detection, even by the ultrasensitive Biacore 2000 system. The background binding (Fig. 7.7) of the loop and enkephalin peptides detected on the

blank flowcell with only the immobilised 9E10, was probably due to the fact that the 9E10 antibody (gift from Dr G.Winter) was not completely pure. In addition, there was insufficient material to warrant purification. For this reason, further dose dependent Biacore experiments were not carried out and efforts focused on improving GlaMor scFv expression and obtaining more material which would then be used in binding assays and quantitative analysis.

9

Discussion and Conclusions

9.1 Gloop2 and GlaMor CDR sequences

The CDR sequences predicted by the computational design (Chapter 3) and the sequences identified by phage selection and used thereafter in bacterial expression studies are summarised in Table 9.1. The mutations found in GlaMor after the phage Elisa's showed highest similarity with the fourth ranked construct out of the ten lowest energy conformations shown in Table 3.3. From this table several conclusions can be drawn:

1. The residues selected by the *ab initio* design for the theoretical antibody combining site are mainly bifunctional residues which are defined as residues capable of more than one type of interaction, e.g Tyr which can engage in hydrophobic packing interactions and can hydrogen bond through its –OH group. These residue

types were selected because the average distance between antibody and antigen is approximately the same as the side chain length (4.2-4.3A) of the selected residue types (His, Phe, Trp, Tyr, Gln, Asn).

2. Residues of the L2 CDR are not included in the design because small antigens tend to bind to those residues located at or near the centre of the combining site. By contrast, larger protein antigens whose interacting surfaces may be as much as 400Å^2 will obligatorily be in contact distance from L2 residues (MacCallum *et al.*, 1996). Notwithstanding this physical proximity, it has been reported that CDR-L2 only rarely contributes to antigen binding (Wilson and Stanfield, 1993).

3. The L3 89 position.

The primer conferring the mutation at this position was degenerate (Table 4.1) encoding for tryptophan, or tyrosine, or a stop codon or cysteine in that position. If a stop codon was introduced this would result in non-viable phage since the transcription/translation mechanism would terminate before the gIII protein. With an additional cysteine the GlaMor scFv construct has an odd number of 5 cysteines, which as discussed later, may have contributed to the low expression yields of the soluble GlaMor scFv.

4. The H1 32 and 34 positions

These two residue positions differed between the computational and the experimental work (noted with asterisks in Table 8.1). The *ab initio* method predicted a F32K mutation and no mutation for position 34 which is an isoleucine in Gloop2 wild type sequence. The PCR mutagenesis primers did not encode for any of these mutations

mutation and no mutation for position 34 which is an isoleucine in Gloop2 wild type sequence. The PCR mutagenesis primers did not encode for any of these mutations that were identified in the construct selected through phage ELISA (Chapter 3). These mutations might be attributable to errors of the DNA polymerase during the PCR cycles. Since isoleucine in position 34 is classified as a canonical residue (Chothia, 1989) its replacement by an arginine may have affected the conformation of H1 and hence the specificity of the GlaMor binding to enkephalin. Database searching indicated that Arginine at this position does not occur in any of the known antibody H1 canonical classes.

Possible interactions between enkephalin and GlaMor are indicated in Fig.9.1

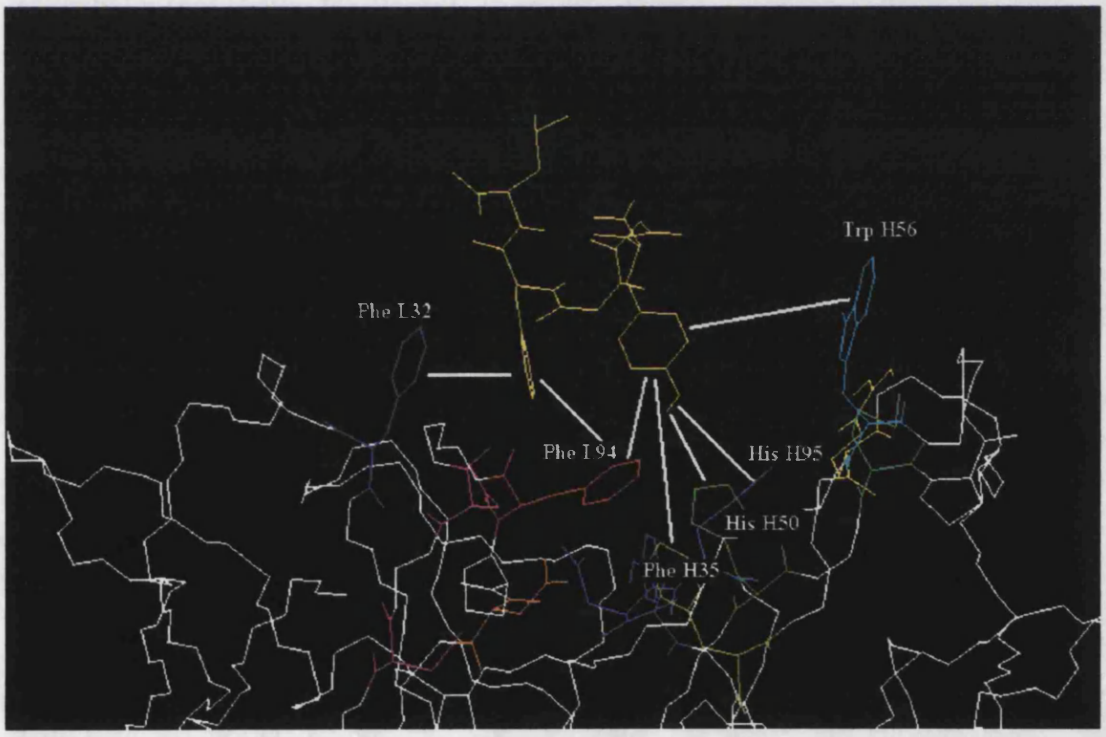


Fig.9.1 As Fig 3.5C but with residues from the 4th design sequence and showing arrows for CDR-enkephalin interactions suggested by the pharmacophore in Fig 3.2

In the orientation generated by the design process a number of plausible pharmacophoric interactions can be satisfied. For example, the tyrosine of enkephalin is flanked by phenylalanine H35 (threonine in Gloop 2), tryptophan H56 (lysine in Gloop 2) with a possible but more distant contact with phenylalanine L94 (tyrosine on Gloop 2). It is also in hydrogen bonding distance from histidine 50 of H2 and histidine 95 of H3, both of which are glutamic acid residues in Gloop2 and possibly either or both involved in the salt bridge interactions with the critical binding arginine of the loop peptide. Additionally, the phenylalanine sidechain of enkephalin is also close to phenylalanine L32 (tyrosine in Gloop 2). More distant, and hence less likely to be important, interactions were seen for D91 and D96 with the backbone of enkephalin (not shown for clarity). These plausible interactions at least are consistent with the requirement (see Fig.3.2) for three hydrophobic sites (H56, L94 and L32) plus an anionic site (H50 and H95). The lack of a clear phenolic site would be an obvious case for second round improvement.

A question that could be posed is : Why was the 4th ranked construct selected and no others? The top ranked, 4th ranked and wild type sequences at the design positions are summarised in Table 9.2 below:

	L1	**L3**	**H1**	**H2**	**H3**
	32	89 91 94 96	32 35	50 56	95 96 101
Top	F	W D F N	K F	H W	H D F
4TH	F	w/C D F D	K F	H W	H D Y
WT	Y	L Y Y L	F T	E K	E I R

Table 9.2

This indicates that there are few differences between the two design sequences – DvN at position L96 and YvF at position H101. The process of modelling and energy calculations is not yet so precise that it can easily discriminate models with this type of subtle difference (although D is charged, N has the ability to hydrogen bond, both as a donor and acceptor). Since there are no positively charged residues in enkephalin the hydrogen bonding of N96 may have been more important than the charge, resulting in deselection of the top candidate.

In contrast, the differences from the Gloop 2 sequence are quite significant. L1 and L3 are essentially hydrophobic in Gloop 2, while the negative charges are located in the heavy chain (H2 and H3). This is essentially reversed in the GlaMor sequences where two histidines are found in the heavy chain, one in place of E50 and the other in place of E95. In addition, two negatively charged residues are located in L1 and L3 essentially altering that region of the groove from hydrophobic to hydrophilic.

A complicating factor was the presence of two unexpected mutations in the selected phage construct. The first at position L89, where a cysteine was present in place of the expected tryptophan, is not likely to have had a major impact since it would probably have been too short to engage in enkephalin contact. Of more consequence could have been the mutation of I34 to arginine in H1. This is a canonical position and is involved in H1 packing. A change in conformation of H1 may thus have occurred. However, position 34 is not a groove contact residue and since there is also a positively charged residue at position 32 (K) it is possible that even with a small conformational change, positive charge at or close to the groove would have been retained.

The foregoing highlights the difficulties of interpreting binding data of mutants in the absence of x-ray structures. Had it been possible to generate accurate models of the selected designs, at least some attempt at rationalisation of the specificity change could have been made. Poor expression thwarted such attempts at full characterisation, and ultimately structural studies. It may be that any redesign project of this sort should only be attempted or pursued where a high chance of obtaining crystals of the mutants exists.

CDR	**L1**	**L3**				**H1**			**H2**		**H3**		
Residue Position Number	32	89	91	94	96	32	34	35	50	55	95	96	101
Gloop2 Sequence	Y	L	Y	Y	L	F	I	T	E	K	E	I	R
Fourthly Ranked Construct	F	W	D	F	D	K	I	F	H	W	H	D	Y
PCR Mutagenesis	F	W/Y/C/ stop	D	F	D	F	I	F	H	W	H	D/N	F/Y
Phage Selected Construct	F	C	D	F	D	F	R	F	H	W	H	D	Y
						*	*						

Table 9.1: Summary table of the CDR sequences predicted computationally and identified experimentally.
The Gloop2 wild type sequence is shown on the first row. Asterisks mark the residue positions were the predicted and the observed sequence differ. The isoleucine at position H34 (marked in grey)was maintained at the design as it is a canonical position. The mutation of this position to arginine as seen on the phage selected construct can attributed to a PCR error.

9.2 Bacterial expression of Gloop2 and GlaMor scFv.

This study defines several problems encountered when periplasmic expression is attempted in *E. coli*, associated mainly with the very low yields of the GlaMor construct. Gloop2 expression levels were also low with 2μg/ml of culture as the maximum concentration obtained. Several optimisation trials have been carried out to maximise the expression yield as described in Chapter 6. The results were consistent for the two constructs with the Gloop2 giving better yields than GlaMor. A lower amount of scFv in the periplasmic or culture supernatant led to less pure material after the IMAC purification. Trials aimed at concentrating the fractions obtained were laborious and extended the time of the experiments, which in turn increased the risk of proteolytic degradation.

Attempts to improve the expression yields of GlaMor included the addition of redox buffer in the media during induction as well as sucrose in a separate set of trials. Both methods were based on earlier studies by Proba *et al.*(1997) and Kipriyanov *et al.* (1997) respectively. The addition of a low concentration redox buffer was attempted to improve the folding of GlaMor scFv, which might have been affected by the fifth, unpaired cysteine introduced during PCR mutagenesis. Normal formation of disulphide bonds between cysteine pairs involves membrane bound proteins that are displaced by the "normal" partners. If the second is missing the unpaired cysteine might form incorrect disulphide bonds with wrong cysteines from other proteins leading to an aggregation-prone product. Glutathione derivatizes the unpaired cysteine, does not interfere with affinity purification and can be removed by mild

reduction conditions. However, these studies showed no effect of glutathione on the expression yields of GlaMor scFv.

Sucrose added in the media has been proposed to stabilise the expressed scFv by inhibiting aggregation. The hypothesis is that sucrose increases the osmotic pressure and causes the periplasm to enlarge (Kiefhaber *et al.*, 1991) decreasing the concentration of secreted protein and resulting in less aggregation. The studies reported here showed that although in the Gloop2 system the addition of sucrose increased the yield of the scFv that was directed into the medium after a 24-hour induction by up to 100%, it had no effect on the expression levels of GlaMor.

It was concluded that the very low efficiency of expression for the GlaMor scFv construct was probably attributable to its highly engineered primary sequence. It has been shown that the degree of successful folding depends on the primary sequence of the variable domains (Plüctkthun, 1994; Knappik and Plückthun, 1995).

Further engineering of the antibody sequence has been proposed as a strategy to obtain correctly folded periplasmic proteins in high yields (Knappik and Plückthun, 1995). Therefore it is possible that further mutagenesis of the GlaMor sequence could improve the expression yields and facilitate further characterisation.

The possibility of producing individually the ten GlaMor constructs as predicted in Chapter 3 using site-specific mutagenesis and further expression and binding analysis would be worth investigating. *In vitro* scanning saturation mutagenesis of an anti-

digoxin antibody binding pocket has been demonstrated as a useful tool for protein structure-function and engineering studies (Burks *et al.*, 1997).

Yeast, insect and mammalian expression systems have been also exploited with relative advantages and disadvantages. The baculovirus mediated expression systems are the most popular of the insect cell expression systems and have already been used for the production of an anti-pancarcinoma antigen scFv and its IL-2 fusion protein (Bei *et al.,* 1995). The main advantage of this system is that can provide high levels of the desired protein with most of the post-translational alterations which are used by higher eukaryotes. However, it has been shown that the nature of the gene to be expressed is able to affect the yields of the expressed protein (Ellis *et al.*, 1988).

The main advantage of yeast over other expression sytems are related to the fact that it both a microorganism and a eukaryot, providing an advanced folding protein pathway for heterologous proteins as well as rapid production on simple growth media. High levels of secreted recombinant antibody scFv's have been achieved in yeast expression systems (Eldin *et al.*, 1997).

Mammalian cells have been also used for the production of a wide range of scFv, scFv fusion proteins in lower levels compared to the yield obtained from the bacterial expression system (Dorai *et al.*, 1994). The main advantage of this system is that the signals for synthesis, processing and secretion of eukaryotic proteins are properly and efficiently recognised.

It should be noted that different antibodies pose their own problems and so no "universal" expression system can be used. Similarly to GlaMor, some antibodies might be inherently poor at being expressed. In such circumstances a choice should be made on investing more time on this particular molecule, in an attempt to improve the expression, or reselecting an alternative molecule that may prove superior. Given the amount of effort and time already expended, in Bath and Vienna, it may be more profitable to reselect an antibody sequence from the original design. This is clearly not an option for the present author.

References

Aman, E., Ochs, B. and Abel, K.J. (1988). Tightly regulated *tac* promoter vectors for the expression of unfused and fused proteins in *Escherichia coli*. *Gene*. **69**, 301-315.

Aman, E. (1985). Plasmid vectors for the regulated, high level expression of eukaryotic genes in *Escherichia coli*. *Developments in Biological Standardization*, **59**, 11-22.

Amit, A.G., Mariuzza, R.A., Phillips, S.E.V. and Poljak, R.J. (1986). 3-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science*, **233**; 747-753.

Arevalo, J.H., Taussig, M.J. and Wilson, I.A. (1993). Molecular basis of crossreactivity and the limits of antibody-antigen complementarity. *Nature* **365**; 859-863.

Artymiuk, P.J. and Blake, C.C. (1981). Refinement of human lysozyme at 1.5 A resolution analysis of non-bonded and hydrogen bond interactions. *J. Mol. Biol.* **152**, 737-762.

Aubry, A., Birlirakis,N., Sakarellos-Daitsiotis, M., Sakarellos, C. and Marraud, M., (1988). Relationship of the crystal and molecular structure of leucine-enkephalin trihydrate to that of morphine. *J. Am.Chem. Soc.*, **C**, 963-964.

Auger, E.A. and Bennet, G.N. (1987). Temperature optimization of *in vivo* expression from the *E. coli trp* and *trp:lac* promoters. *Biotechnol. Lett.*, **8**, 157-162.

Bagshawe, K.D. (1995). Antibody directed enzyme prodrug therapy: A review. *Drug Dev. Res.* **34**; 220-230.

Balass, M., Heldman, Balass, M., Heldman, Y., Cabilly, S., Givol, D., Katchalski-Katzir, E. and Fushs, S. (1993). Identification of a hexapeptide that mimics a conformation-dependent binding site of acetylcholine receptor by use of a phage-display library. *Proc. Natl. Acad. Sci. USA* **90**; 10638-10642.

Balbas, R., and Bolivar, F. (1990). Design and construction of sxpression plasmid vectors in *Escherichia coli*. *Methods Enzymol.* **185,** 14-37.

Balint, R.F. and Larrick, J.W. (1993). Antibody engineering by parsimonious mutagenesis . *Gene*, **137**, 109-118.

Bantley, G.A., Boulot, G., Riottot, M.M. and Poljak, R.J. (1990). 3-dimensional structure of an idiotope-anti-idiotope complex. *Nature*, **348**, 254-257.

Barbas III, C.F., Kang, A.S., Lerner, R.A. and Bencovic, S.J. (1991). Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc. Natl. Acad. Sci. USA* **88**; 7978-7982.

Barettino, D., Feigenbuitz, M., Valcarcel, R. and Stunnenberg, H.G. (1994). Improved method for PCR-mediated site-directed mutagenesis. *Nuc. Ac. Res.* **22**; 541-542.

Basak, A., Jean, F., Dugas, H. and Lazure, C. (1994). Biotinylation of an enkephalin-containing heptapeptide via various spacer arms. Synthesis, comparative binding studies towards avidin, and application as substrates in enzymatic reactions. *Bioconjugate Chem.* **5**; 301-305.

Beckwith, J.R. (1987). The lactose operon. In: *Escherichia coli* and *Salmonella typhimurium. Cellular and Molecular Biology* (Neidhardt FC, JL Ingraham, KB Low, B Magasanik, M Schaecheter and HE Umbarger, Eds.). American Society for Microbiology, **2**, 1444-1452, Washington, DC.

Bei, R., Schlom, ., Kashmiri, S.V.(1995). Baculovirus expression of a functional singl-chain immunoglobulin and its IL-2 fusion protein. *J. Immunol. Methods*, **186**, 245-250.

Bentley, G.A., Bhat, T.N., Boulot, G., Fischmann, T., Navaza, J., Poljak, R.J., Riottot, M.M. and Tello, D. (1989). *Cold Spring Harbor Symposium on Immune Recognition.*

Bentley, G.A., Boulot, G., Riottot, M.M. and Poljak, R.J., (1990). 3-Dimentional structure of an idiotope anti-idiotope complex *Nature* **348**, 254-257.

Berek, C. and Milstein, C. (1987). Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.* **96**; 23-41

Better, M., Chang, C.P., and Robinson, R.R. (1988). *Escherichia coli* secretion of active chimeric antibody fragment. *Science* **240**; 1041-1046

Benvenuto, E., Pardas, R.J., Tavazza, R., Ancora, G., Biocca, S., Cattaneo, A. and Galeffi, P. (1991) Phytoantibodies: a general vector for the expression of immunoglobulin domains in transgenic plants. *Plant Mol. Biol.* **17**:865-874.

Bhat, T.N., Bentley, G.A., Boulot, G., Greene, M.I., Tello, D., Dall'Acqua, W., Souchon, H., Schwarz, F.P., Mariuzza, R.A. and Poljak, R.J. (1994) Bound water molecules and conformational stabilisation help mediate an antigen-antibody association. *Proc. Natl. Acad. Sci. USA* **91** 1089-1093.

Bird, R.E., Hardman, K.D., Jacobson, J.W., Jonson, S., Kaufman, B.M., Lee, S.M., Lee, T., Pope, S.H., Riordan, G.S. and Whitlow, M. (1988). Single-chain antigen-binding proteins. *Science* **242**; 423-431.

Bird, R.E. and Walker, B.W. (1991). Single chain antibody variable regions. *TIBTECH* **9**; 132-140.

Blake, C. C.F., Koenig, D.F., Mair, G.A., North, A.C.T., Phillips, D.C. and Sarma, V.R. (1965). Structure of hen egg-white lysozyme. A three-dimentional fourier synthesis at 2 A resolution *Nature* **206**; 757-761.

Bole, D.G., Henderhot, L.M and Kearney, J.F. (1986). Post-translational association of immunoglobulin heavy chain binding protein with nascent heavy chain in non-secreting hybridomas. *J. Cell Biology* **102**; 1558-1566.

Borrebaeck, C.A.K., Malborg, A.C., Furebring, C., Michaelsson, A., Ward, S., Danielsson, L. and Onlin, M. (1992). Kinetic analysis of recombinant antibody-antigen interactions. Relation between structural domains and antigen binding. *Biotechnology*, **10**, 687-698.

Botha, S.A. and Flanagan, D.R. (1992). Characterization of digoxin hydrates using thermal methods. *Int. Jour. Pharmaceutics* **82**; 185-194.

Botha, S.A. and Flanagan, D.R. (1992). Non-thermal methods in characterization of anhydrous digoxin and two digoxin hydrates. *Int. Jour. Pharmaceutics* **82**; 195-204.

Bott, R. and Sarma, R. (1976). Crystal structure of turkey egg-white lysozyme. Results of the molecular replacement method and 5Å resolution. *J. Mol. Biol.* **106**, 1037-1046.

Bottger,V. and Birgitte Lane, E. (1994). A monoclonal antibody epitope on Kerratine 8 identified using a phage peptide library. *J. Mol.Biol.* **235**; 61-67.

Breitling, F., Dubel, S., Seehaus,T., Klewinghaus, I. and Little, M. (1991). A surface expression vector for antibody screening. *Gene* **104**; 147-153.

Buchner, J. and Rudoloph, R. (1991). Renaturation, purification and characterisation of recombinant Fab fragments produced in *Escherichia coli. Biotechnology* **9**; 157-164

Burks E.A. Chen, G., Georgiou, G. and Iverson, B.L. (1997). *In vitro* scanning saturation mutagenesis of an antibody binding pocket. *Proc. Natl Acad. Sci USA* **94**, 412-417

Bullock, W.O. (1987). XL1-Blue: A high efficiency plasmid transforming RECA Escherichia coli strain with β-galactosidase selection., *Biotechniques*, **5**, 376.

Burton, D.R. (1991). Human and mouse monoclonal antibodies by repertoire cloning. *Trends in Biotechnology* **9**; 169-175.

Bye, E., (1976). The crystal structure of morphine hydrate. *Acta Chemica Scandinavica*, **30**, 549-554.

Carter, P., Kelley, R.F., Rodrigues, M.L., Snedecor, B., Covarrubias, M., Velligan, M.D., Wong, W.L.T., Rowland, A.M., Kotts, C.E., Carver, M.E., Yang, M., Bourell, J.H., Shepard, H.M. and Henner, D. (1992). High-level *Exscherichia coli* expression

and production of a bivalent humanised antibody fragment. *Bio/Technology,* **10,** 163-167.

Clackson, T., Hoogenboom, H.R., Griffiths, A.D. and Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature* 352; 624-627.

Chalmers, J.J., Kim, E., Telford, J.N., Wong, E.Y., Tacon, W.C., Shuler, M.L. and Wilson, D.B., (1990). Effects of temperature on E. coli overproducing b- lactamase of human epidermal growth factor. *Applied and environmental microbiology.* **56,** 104-111.

Cheettham, J.C., Raleigh, D.P., Griest, R.E., Redfield, C., Dobson, C.M. and Rees, A.R. (1991). Antigen mobility in the combining site of an anti-peptide antibody. *Proc. Natl. Acad. Sci. USA* **88**; 7968-7972.

de la Cruz, X., Mark, A.E., Tormo, J., Fita, I. and Vangusteren, W.F. (1994). Investigation of shape variations in the antibody binding site of molecular dynamics computer simulation. *J. Mol. Biol.* **236**;1186-1195.

Chiswell, D.J. and McCafferty, J. (1992). Phage antibodies;will new "coliclonal" antibodies replace monoclonal antibodies? **10**; 80-89.

Chothia, C. Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E., Davies, D., Tulip, W.R., Colman, P.M., Spinelli, S., Alzri, P.M. and Poljak, R.J. (1989). Conformations of immunoglobulin hypervariable regions. *Nature* **342**; 877-883.

Chothia, C. (1991). Antigen recognition. *Cur. Op. Str. Biol.* **1**; 53-59.

Chothia, C. and Lesk, A.M., (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.*, **196**, No.4, 901-917.

Chothia, C., Lesk, A.M, Levitt, M., Amit, A.G, Mariuzza, R.A, Phillips, S.E.V, Poljak, R.J (1986). The predicted structure of immunoglobulin-d1.3 and its comparison with the crystal-structure. Science, **233**, No.4765,.755-758

Colman, P.M. (1994). Effects of amino-acid sequence changes on antibody-antigen interactions. *Research in Immunology*, **145**, No 1, 33-36.

Colcher, D., Bird, R. and Roselli, M. (1990). *In vivo* tumour targeting of a recombinant single-chain antigen-binding protein. *J. Natl. Cancer Institute*, **82**; 1191-1197.

Covell, D.G. and Mallqvist, A. (1997). Analysis of protein-protein interactions and the effects of amino acid mutattions on their energetics. The importance of water molecules in the binding epitope., *J. of Mol. Biol.*, **269**, No.2, 281-297.

Darsley, M.J. and Rees, A.R, (1985a). 3 distinct epitopes within the loop region of hen egg lysozyme defined with monoclonal-antibodies. *EMBO J.*, **4**, 383-392.

David, D. (1990). Section II: Expression in Escherichia coli. *Methods in Enzymology,* **185,** 3: Gene expression technology. San Diego, Academic Press, Inc.

Darsley, M.J. and Rees, A.R, (1985b). Nucleotide sequences of 5 anti-lysozyme monoclonal antibodies. *EMBO J.* **4,** 393-398.

Darveau, R.P., Somerville, J.E and Fell, P.H. (1992). Expression of antibody fragments in *Escherichia coli. J. Clinical Immunoassay.* **15** No. 1; 25-29.

Daugherty B.L., DeMartino, J.A., Law, M.F., Kawka,D.W., Singer, I.I. and Mark, G.E. (1991). Polymerase chain reaction facilitates the cloning, CDR-grafting, and rapid expression of a murine monoclonal antibody directed against the CD18 component of leucocyte integrins. *Nuc. Ac. Res.***19** No. 9; 2471-2476.

Davies, D.R., Sheriff, S. and Padlan, E.A. (1988). Antibody- antigen complexes. *J. Biol. Chem.* **263** No. 22; 10541-10544.

Davies, D.R. and Padlan, E.A. (1990). Antibody-antigen complexes. *Annu. Rev. Biochem.* **59**; 439-473.

de la Paz, P., Sutton, B., Darsley, M.J. and Rees, A.R. (1986). Modelling of the combining sites of three anti-lysozyme monoclonal antibodies and of the complex between of the antibodies and its epitope. *EMBO J.*, **5**, 415-425.

Deev, S.M. and Polyanovskii (1995). Antibody engineering (A Review). *Molecular Biology* **28** No.6; 789-795.

Dente, L., Cesareni, G., Micheli, G., Felici, F., Folgori, A., Luggago, A., Monaci, A. and Delmastro, P. (1994). Monoclonal antibodies that recognise filamentous phage: tools for phage display technology. *Gene* **148**; 7-13.

Deonarain, M.P., Rowlinson-Busza, G., George, A.J.T. and Epenetos, A. A. (1997). Redesigned anti-human placental alkaline phosphatase single-chain Fv: soluble expression, characterization and *in vivo* tumour targeting. *Prot. Engineering*, **10**, No. 1, 89-98.

Devlin, J.J., Panganiban, L.C. and Devlin, P. (1990). Random peptide libraries: a molecular source of specific protein binding molecules. *Science* **249**; 404-406.

Dorai, H., McCartney, J.E. and Hudziak, R.M. (1994). Mammalian cell expression of single-chain Fv (scFv) antibody proteins and their C-terminal fusions with interleukin-2 and other effector domains. *BioTechnology*, **12**, 890-899.

Dubel, S., Breitling, F., Fuchs, P., Braunagel, M., Klewinghaus, I. and Little, M. (1993). A family of vectors for surface display and production of antibodies. *Gene* **128**; 97-101.

Edelman, G. M. (1970). *Biochemistry,* **9**, 3167-3205.

Eldin, P., Pauza, M.E. and Heida, Y.(1997). High level secretion of two antibody single-chain Fv fragments by *Pichia pastoris. J. Immunol.Methods,* **201**, 67-75,

Ellis, L., Levitan, A., Cobb, M.H and Ramos, P. (1988). Efficient expression in insect cells of a soluble, active human insulin receptor protein-tyrosine kinase domain by use of a baculovirus vector. *J. Virol.* , **62**, 1634-1640.

Evan G.I., Lewis, G.K., Ramsay, G. and Bishop, J.M. (1985). Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. and Cell. Biol.*, **5**, No. 12, 3610-3616.

Field, H., Yarranton, G.T. and Rees, A.R. (1989). Expression of mouse immunoglobulin light and heavy chain variable regions in *Escherichia coli* and reconstitution of antigen binding activity. *Protein Engineering* **3** No.7; 641-647.

Field, H. (1988): *D. Phil.* Thesis, Oxford.

Foote, J. and Eisen, H.N. (1995). Kinetic affinity limits on antibodies produced during immune responses. *Proc. Natl. Acad. Sci. USA* **92**; 1254-1256.

Froyen, G., Ronsse, I. and Billiau, A. (1993). Bacterial expression of a single-chain antibody fragment (scFv) that neutralises the biological activity of human interferon-γ. *Molecular Immunology* **30** No. 9; 805-812.

Garrard, L.J. and Henner, D.J (1993). Selection of an anti-IGF-1 Fab from a Fab phage display created by mutagenesis of multiple CDR loops. *Gene* **128**; 103-109.
Garrard, L.J. and Zhukovsky, E.A. (1992). Antibody expression in bacteriophage systems: the future of monoclonal antibodies. *Cur.Op. Str. Biol.* **3**; 474-480.

Gibson, T.J. (1984). *Ph.D* Thesis, Cambridge University.

Gilliland, L.K., Norris, N.A., Marquardt, H., Tsu, T.T., Hayden, M.S., Neubauer, M.G., Yelton, D.E., Mittler, R.S. and Ledbetter, J.A. (1996). Rapid and reliable cloning of antibody variable regions and Generation of recombinant single chain antibody fragments. *Tissue Antigens* **47**; 1-20.

Glockshuber, R., Malia, M., and Pfinginger, I. (1990) A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochemistry* **29**; 1362-1371.

Glockshuber, R., Scmidt, T. and Pluckthun A. (1992) The disulpfide bonds in antibody variable domains: Effects on stability, folding *in vitro*, and functional expression in *Escherichia coli. Biochemistry* **31**, 1270-1276

Gold, L., (1988). Posttranscriptional regulatory mechanisms in *Escherichia coli. Annu. Rev. Biochem.* **57**, 199-233.

Goodsell, D.S., Olson, A.J. (1990). Automated docking of substrates to proteins by simulated annealing. *Prot. Struct. Fun. and Gen.*, **8**, No.3, 195-202.

Griffiths, A.D., Williams, S.C., Hartley, O., Tomlinson, I.M., Waterhouse, P., Grosby, W.L., Kontermann, R.E., Jones, P.T., Low, N.M., Allison, T.J., Prospero, T.D., Hoogenboon, H.R., Nissim, A., Cox, J.P.L.C. Harrison, J.L., Jaccolo, M., Gherardi., E. and Winter, G. (1994). Isolation of high affinity human antibodies directly from large synthetic repertoires. *The EMBO J.*, **13**, 3245-3260.

Griffiths, A.D., Malmqvist, M., Marks, J.D., Bye, J.M., Embleton, M.J., McCafferty, J., Baier, M., Holliger, K.P., Gorick, B.D., Hughes-Jones, N.C., Hoogenboom, H.R. and Winter, G. (1993). Human anti-self antibodies with high specificity from phage display libraries *EMBO J.* **12** No. 2; 725-734.

Gussow, D., Ward, E.S., Griffiths, A.D., Jones, P.T., and Winter, G. (1989) Generation of binding activities from *Escherichia coli* by expression of a repertoire of immunoglobulin variable domains. *Cold Spring Harb. Symp. Quant. Biol.* **1**; 265-273.

Hanahan, D. (1983). Studies of transformation of *Escherichia coli* with plasmids. J. of *Mol. Biol.* **166**, 557-580.

Hatt, J., Callahan, M., and Greener, A. (1992). *Strategies*, **5**, 2-3.

Hawkins, R.E., Russell, S.J. and Winter, G. (1992). Selection of phage antibodies by binding affinity; mimicking affinity maturation. *J. Mol. Biol.* **226**; 889-896.

He, M.Y., Hamon, M., Liu, H., Kang, A. and Taussing, A.J. (1995). Functional expression of an anti-progesterone antibody fragment in the cytoplasm of a mutant *Escherichia coli. Nuc. Acid. Res.***23**; No.19, 4009-4010.

Herron, J.N., He, X.M., Ballard, D.W., Blier, P.R., Pace, P.E., Bothwell, A.L.M., Voss, E.W. and Edmundson, A.B., (1991). An autoantibody to single-stranded-DNA; Comparison of the 3-dimensional structures of the unliganded Fab and a deoxynucleotide Fab complex. *Prot. Struct. Funct. and Genet.*, **11**, No.3, 159-175.

Herron, J.N., He, X.M., Mason, M.L., Voss, E.W. and Edmundson, A.B. (1989). 3-dimensional structure of a fluorescein Fab complex crystallised in 2-methyl-2,4-pentanediol. *Prot. Str. Func. Gen.* **5**, 271-280.

Hilyard, K.L., Staunton, D., Jones, A.E and Rees, A.R. (1992*) Protein Enginering of antibody combining sites.* A practical appproach, IRL press.

Hochuli, E., Dobeli, H. and Schacher, A. (1988). New metal chelate absorbent selcetive for proteins and peptides containing neighbouring histidine residues. *J. of Chromatography*, **411**, 177-184.

Hoogenboom, H.R.(1997). Designing and optimizing libarry selection strategies for generating high affinity antibodies. *Trends in Biotechnology.* **15**, 62-70.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P. and Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nuc. Acid Res.* **19** No.15; 4133-4137.

Hoogenboom, H.R., Marks, J.D., Griffiths, A.D. and Winter, G. (1992). Building antibodies from their genes. *Immunological reviews* **130**; 41-67.

Hoogenboom, H.R. and Winter, G. (1992). By-passing immunisation; Human antibodies from synthetic repertoires of germline VH Gene segments rearranged *in vitro*. *J. Mol. Biol.* **227**; 381-388.

Huse, W.D., Stinchcombe, T.J., Glaser, S.M., Starr, L., Maclean, M., Hellstrom, K.E., Hellstrom, I. And Yelton, D.E. (1992). Application of a filamentous phage PVIII fusion protein system suitable for efficient production, screening and mutagenesis of Fab antibody fragments. *J. Immunology*, **149**, 3914-3920.

Huse, W.D., Sastry, L., Iverson, S.A., Kang, A.S., Alting-Mees, M., Durton, D.R., Bencovic, S.J. and Lerner, R.A. (1989). Generatin of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* **246**; 1275-1281.

Huston, J.S., Levinson, D., Mudgett-Hunter, H.M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R. and Opperman, H. (1988). Protein engineering of antibody-binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **85**; 5879-5888

Huston, J.S., Mudgett-Hunter, H.M., Mei-Sheng, T., McCartney, J., Warren, F., Haber, E. and Opperman, H. (1991). Protein engineering of single-chain Fv analogs and fusion proteins. *Meth. Enzymol.* **203**; 46-55

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M-S, Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R. and Oppermann, H. (1988). Protein engineering of antibody combining sites recovery of specific activity in an anti-digoxin single chain Fv analog produced in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **83**, 5879-5883.

Ingraham, J. (1987). Effect of temperature, pH, water, activity and pressure on growth. In: *Escherichia coli* and *Salmonella typhimutium*. *Cell. And Mol. Biol.* (Nedhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B. Schaechter, B. amd Umbarger, eds.). **2**, 1543-1554, *American Society for Microbiology*; Washington DC.

Jeffrey, P. (1989).*D. Phil.* Thesis, Oxford.

Johnsson, B., Lofas, S. and Lindquist, G. (1991). Immobilisation of proteins to carboxylmethyl dextran- modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal. Biochem.*, **198**, 268-277.

Johnson, K.S. and Chiswell, D.J. (1993). Human antibody engineering. *Cur. Op. Str. Biol.* **3**; 564-571.

Jones, S.T., Dear, P.H., Foote, J., Neuberger, M.S. and Winter, G. (1986). Replacing the complementarity determining regions in a human antibody with those from a mouse. *Nature* 321; 522-534.

Jung, S. and Pluckthun, A. (1997). Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting Protein Eng. Vol.10, 959-966

Kabat, E., Wu, T.T., Reid-Meiller, M., Perry, H.M., Gottesman, K.S. and Foeller, C. (1987). Sequences of proteins of immunological interest. *US Department of Health and Human Services, U.S.A* Government Printing Office.

Kang, A.S., Barbas, C.F., Janda, K.D., Benkovic, S.J. and Lerner, R.A. (1991). Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA.* **88,** 4363-4370.

Karle, I.L., Flippenanderson,J.L., Uma, K. and Balaram, O., (1990). Apolar peptide models for conformational hetero*Gene*ity, hydration, and packing of polypeptide halices-crystal-structure of heptapeptides and octapeptides containing alpha-aminoisobutyric acid. Proteins: *Struct., Funct., Genet.,* 7 No. 1, 62-73.

Karlsson, R., Michaelsson, A. and Mattson, L. (1991). Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system. *J. Immunol. Methods*, **145**, 229-240.

Kennard, O., (1991). The Cambridge crystallographic databank. Crystal structure data for about 90.000 organic and organo-metallic compounds. *Cambridge Crystallographic Data Centre*, Released bi-annual in January and July.

Kerschbaumer, R.J., Hirschl, S., Kaufmann,A., Ibl, M., Koenig, R. and Himmler, G. (1997). Single-chain Fv fusion proteins suitable as coating and detecting reagents in a double antibody sandwich enzyme-linked immunosorbent assay. *Anal. Biochem.,* **249,** 219-227.

Kerschbaumer, R.J., Hirschl, S., Schawger,C., Ibl, M. and Himmler, G. (1996). pDAP2: a vector for constructionof alkaline phosphatase fusion-proteins. *Immunotechnology,* **2,** 145-150.

Kiefhaber, T., Rudolph, R., Kohler, H.H. and Buchner, J. (1991). Protein aggregation of *in vitro* and *in vivo*; a quantitative model of the kinetic competition between folding and aggregation. *Biotechnology*, **9**, 825-829.

Kipriyanov, S.M., Moldenhauer, G and Little, M. (1997a). High level production of soluble single chain antibodies in small-scale *Escherichia coli* cultures. *J. of Immun. Methods.,* **200**, 69-77.

Kipriyanov, S.M., Moldenhauer, G., Martin, A.C.R, Kupriyanova, O.A. and Little, M. (1997b). Two amino acid mutations in an anti-human CD3 single chain Fv antibody fragment that affect the yield on bacterial secretion but not the affinity. *Prot. Engineering*, **10**, No.4, 445-453.

Knappik, A. and Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its *in vivo* folding. *Protein Engineering*; **8** No.1; 81-89.

Kohler, G. and Milstein, C., (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature (Lond.).* **256**, 495-497.

Krebber, C., Spada, S., Desplancq, D. and Pluckthun, A., (1995). Co-selection of cognate antibody-antigen pirs by selectivly infective phages. *FEBS Letters*, **377**, 227-231.

Krook. M., Mosbach, K. and Lindbladh, C. (1994). Selection of peptides with affinity for single stranded DNA using a phage display library. *Bioch. And Bioph. Res. Commun.* **204** No. 2; 849-854.

Laemmli,U.K (1970). Cleavage of structural proteins during the assembly of the head of bactreriophage T4. *Nature*, **227**, 680-685.

Lei, S.P., Lin, H.C., Wang, S.S., Callaway, J. and Wilcox, G., (1987). Characterization of the *Erwinia carotovora pel*B Gene and its product pectase lyase. *J. Bacteriol.* **169**; 4379-4383.

Li, X., Robbins, J.W. and Taylor, K.B. (1990). The production of recombinant β-galactosidase in *E. coli* in yeast enriched growth medium. *J. Ind. Microbiol.*, **5**, 85-94.

Lonberg, N., Taylor, L.D, Harding, F.A., Trounstine, M., Higgins, K.M., Schramm, S.R., Kuo, C.C., Mashayekh, R., Wymore, K., McCabe, J.G., Munoz-O'Regan, D., O'Donnell, S.L., Lapachet, E.S.G., Bengoechea, T., Fishwild, D.M., Carmack, C.E., Kay, R.M. and Huszar, D. (1994). Antigen-specific human antibodies from mice comprising four distinct Genetic modifications. *Nature* **368**; 856-859.

MacCallum, R.M., Martin, A.C.R. and Thornton, J.M. (1996). Antibody-antigen interactions: contact analysis and binding site topography. *J. Mol. Biol.* **262**; 732-745.
Makowski, L. (1993). Structural constraints on the display of foreign peptides on filamentous bacteriophages. *Gene* **128**; 5-11.

Makrides, S.C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli. Microbiol. Reviews,* **60,** 512-538.

Mamalaki, A., Trakas, N. and Tzartos, S.J. (1993). Bacterial expression of a single-chain Fv antibody fragment which efficiently protects the acetylcholine receptor against antigenic modulation caused by myasthenic antibodies. *Eur. J. Immunol.* **23**; 1839-1845.

Marion, T.N., Tillman, D.M, Jou, N.T. and Hill, R.J., (1992). Selection of immunoglobulin variable regions in autoimmunity to DNA. *Immunol. Reviews*, **128**, 123-149

Mariuzza, R.A.and Poljak, R.J., (1993). The basics of binding mechanisms of antigen recognition and mimicry by antibodies. *Cur. Op. Immunol.*, **5**, No.1, 50-55.

Martin, A.C.R. (1990). *D. Phil Thesis*, Oxford.

Marks, J.D., Griffiths, A.D., Malmqvist, M., Clackson, T.P., Bye, J.M. and Wintern G. (1992). By-passing immunization: Building high affinity human antibodies by chain suffling. *Bio/Technology* **10**; 779-783.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D., and Wintern G. (1991). By-passing immunization; Human antibodies from V-*Gene* libraries displayed on phage. *J. Mol. Biol.* **222**; 581-597.

Martin, A.C.R. (1990). *D. Phil. Thesis*, Oxford.

Martin, A.C.R., Cheetham, J.C. and Rees, A.R. (1989). Modelling antibody hypervariable loops: A combined algorithm. *Proc. Natl. Acad. Sci. USA* **86**; 9268-9272.

Matthews, B.W. (1991). Structure and stability of mutant proteins. *Am. J. of Human* Genetics, **49** No.4, 50-54.

McCafferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990). Phage antibodies: filamentous displaying antibody variable domains. **348**; 552-554.

Milstein, C. (1989). The croonian lecture, Antibodies: a paradigm for the biology of molecular recognition. *Proc. R. Soc. Lond.* B **239**; 1-16.

Michelsen, B.K (1995). Transformation of Escherichia coli increases 260 fold upon inactivation of T4 DNA ligase. *Anal. Biochem.* **225**; 172.

Munro, S. amd Pelham, H.R.B., (1986). An Hsp70-like protein in ER: identity with the 78kDa glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell*, **46**, 291-300.

Nancib, N., Branlant, C. and Boudrant, J., (1991). Metabolic roles of peptone and yeast extract for the culture of a recombinant strain of *Escherichia coli. J. Ind. Microbiol.*, **8**, No.3, 165-170.

Navarro-Teulon, I., Peraldi-Roux, S., Bernardi, T., Marin, M., Piechaczyk, M., Shire, D., Pau, B. and Piechaczyk-Biard, M. (1995). Expression in *Escherichia coli* of soluble and M13 phage-displayed forms of a single-chainantibody fragment specific for digoxin: assessment in a novel drug immunoassay. *Immunotechnology* **1**; 41-52.
Neumann and Ulam, 1945.

Nieba, L., Honegger, A., Krebber, C. and Pluckthun, A. (1997). Disrupting the hydrophobic pathes at the antibody variable/constant domain interface: improved *in vivo* folding and physical characterization of an engineered scFv fragment. *Prot. Engineering,* **10**, 435-444.

Nikaido, H. and Rosenberg, E.Y. (1981): Effect of the solute size on diffusion rates through the transmembrane pools of the outer membrane of *E. coli.* J. Gen.Physiol; **77** 121-127

Nissim, A., Hoogenboom, H.R., Tomlinson, I.A., Flynn, G., Midgley, C., Lane, D. and Winter, G. (1994). Antibody fragments from a "single pot" phage display library as immunochemical reagents. The *EMBO J.* **13** No.3; 692-698.

Novotny, J., (1991). Protein antigenicity: A thermodynamic approach. *Mol.Immunol.*, **28** No 3, 201-207.
Novotny, J.and Sharp, K., (1992). Electrostatic fields in antibodies and antibody antigen complexes. *Progress in biophysics and molecular biology.* **58**, No.3, 203-224.

Orlandi, R., Gussow, D.H., Jones, P.T. and Winter, G. (1989). Cloning immunological variable domains for expressin by the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* **86**; 3833-3837.

Owen, M., Gandecha, A., Cocjburn, B. and Whitelam, G. (1992) Synthesis of a functional anti-phytochrome single chain Fv proetin in transgenic tobacco. *Bio/Technology* **10**; 790-797.

Padlan, E., (1990). On the *Nature* of antibody combining sites: Unusual structural features that may confer on these sites an enhanced capacity for binding ligands. *Proteins: Struct., Funct., Genet.*, **7**, 112-124.

Padlan, E., Silverton, E., Sheriff, S., Cohen, G., Smith-Gill, S. and Davies, D. (1989). Structure of antibody-antigen complex: crystal structure of the HyHEL-10 fab-lysozyme complex. *Proc. Natl. Acad. Sci. USA*, **86**, 5938-5942.

Pantoliano, M.W., Bird, R.E., Johnson, S., Asel, E.D., Dodd, S.W., Wood, J.F and Hardman, K.D. (1991). Conformational stability, folding and ligand binding affinity of single chain Fv immunoglobulin fragments expressed in *Escherichia coli. Biochemistry*, **30**, 10117-10125.

Pedersen, J. (1993). *Ph.D* Thesis, Bath.

Pedersen, J., Searle, S., Henry, A. and Rees, A.R. (1992). Antibody Modeling: Beyond homology. *Immunomethods* **1**; 126-136.

Persson, M.A.A, Caothien, R.H, Burton, D.R. (1991). Generation of diverse high affinities human monoclonal antibodies by repertoire cloning. *Proc. Natl. Acad. Sci. USA*, **88**, 2432-2436.

Poljak, R.J., Amzel, L.M., Avey, H.P., Chen, B.L., Phizackerley, R.P. and Saul, F., (1973). Three-dimensional structure of the Fab' fragment of a human immunoglobulin at 2.8A resolution. *Proc. Nat. Acad. Sci.* (Wash.). **70**, 3305-3310.

Polymenis, M. and Stollar, B.D. (1995). Domain Interactions and antigen binding of recombinant anti-Z-DNA antibody variable domains. *J. Immunol.* **154**; 2198-2208.

Plückthun, A. (1996). Antibody engineering, A practical approach. *IRL Press.*

Plückthun, A. (1992). Mono- and bivalent antibody fragments produced in *Escherichia coli:* Engineering, folding and antigen binding. *Immunological Reviews* **130**; 151-188

Plückthun, A. (1991). Antibody engineering: Advances from the use of *Escherichia coli* expression systems. *Bio/technology* **9**; 545-551.

Porter, R., (1958). Separation and isolation of fractions of rabbit gamma-globulin containing the antibody and antigenic combining site. *Nature*, **182**, 670-671

Proba, K., Honegger, A. and Plückthun, A. (1997) A natural antibody missing a cysteine in VH: Consequences for thermodynamic stability and folding. *J. Mol. Biol.* **265**; 161-172.

Proba, K, Ge, L.M and Pluckthun, A. (1995). Functional antibody single-chain fragments from the cytoplasm of escherichia-coli - influence of thioredoxin reductase (trxb)

Rees A.R., Stauton, D., Webster, D.M., Searle, S.J., Henry, A.H. and Pedersen, J.T. (1994). Antibody design: Beyond the natural limits. *TIBTECH*, **12**, 199-206.

Riechmann, L., Clark, M., Waldmann, H. and Winter, G. (1988) Reshaping human antibodies for therapy. *Nature* **332**; 323-330

Riechmann, L. and Weill (1993). Phage display and selection of a site-directed randomized single-chain antibody Fv fragment for its affinity improvement. *Biochemistry* **32**; 8848-8855.

Rini, J.M., Schulzegahmen, U. and Wilson, I.A., (1992). Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science*, **255**, No 5047, 959-965.

Roberts, S.R. and Rees, A.R. *Protein Eng.* (1986). Roberts, S.R. and Rees, A.R. Protein Eng. The cloning and expression of an antipeptide antibody - a system for rapid analysis of the binding-properties of engineered antibodies, **1**, 59-65.

Roberts, S., Cheetham, J. and Rees, A.R. (1987). *Gener*ation of an antibody with enhanced specificity and affinity for its antigen by preotein engineering. *Nature*, **328**, 731-734.

Roguska, M.A., Pedersen, J.T., Henry, A.H., Roja, C.M., Avery, B., Hoffee, M., Cook, S., Lambert, J.M., Rees, A.R., and Guild, B.C. (1997). A comparison of two murine monoclonal antibodies humanized by CDR-grafting and variable domain resurfacing. *Protein Engineering*, **10**, 2, 181.
Roitt, I. (1984). In *Essential Immunology*, 5th ed. Blackwells Scientific Publications.

Rudikoff, S., Satow, Y., Padla, E., Davies, D. and Potter, M. (1981). Kappa-chain structure from a crystallised murine Fab'. Role of joinng segment in hapten binding. *Mol. Immunol.* **18**, 705-711.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning, A laboratory manual, *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, N.Y.

Samuelsson, E., Johasson, P., Viklund, Nilsson, B. and Uhlen, M. (1996) Affinity-assisted *in vivo* folding of a secreted human peptide hormone in *Escherichia coli. Nature Biotechnology* **14**; 751-755.

Sanger, F., Nicklens, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acd. Sci.* USA **74**, 5463-5467

Schier, R., McCall, L.A., Marshall, K.W., Merritt, H., Yim, M., Crawford, R.S., Weiner, L.M., Marks, C and Marks, J.D. (1996). Isolation of picomolar affinity anti-C-erb-β- 2 single chain Fv by molecular evolution of the complementarity determining regions in the center of the antibody comboning site. *J. Mol. Biol.* **263**, 4, 551-567.

Sherrif, S., Silverton, E.W., Padla, E.A., Cohen, G.H., Smith-Gill, S.J., Finzel, B.C. and Davies, D.R., (1987). Three dimensional structure of an antibody-antigen complex. *Proc. Natl. Acad. Sci. USA*, **84**, 8075-8079.

Shibui, T., Nagahari, K. (1992). Secretion of a functional fab fragment in *Escherichia coli* and the influence of culture conditions. *Appl. Microbiol. and Biotech.*, **37**, No.3, 352-357.

Shoham, M., (1993). Crystal-structure of an anticholera toxin peptide complex at 2.3A *J. of Mol. Biol.*, **232**, No.4, 1169-1175.

Shusta, EV, Raines, RT, Pluckthun, A, Wittrup, K.D. (1998). Increasing the secretory capacity of Saccharomyces cerevisiae for production of single-chain antibody fragments. Nature Biotechnology,.**16**, No.8, pp.773-777

Skerra, A. (1994). Use of the tetracycline promoter for the tight regulated production of a murine antibody fragment in *Escherichia coli. Gene,* **151,** 131-135.

Skerra, A., Pfitzinger, I. and Plückthun, A. (1991). The functional expression of antibody Fv fragments in *Escherichia coli*: improved vectors and a Generally applicable technique. *Bio/Technology* **9**; 273-277.

Skerra, A. and Plückthun, A. (1988). Assembly of a functional immunoglobulin Fv fragmnet in *Escherichia coli. Science*, **240**; 1038-1045.

Smith-Gill, S.J., Mainhart, C., Lavoie, T.B., Feldmann, R.J., Drohan, W. and Brooks, B.R. (1987). A 3- dimensional model of an anti-lysozyme antibody. *J. Mol. Biol.* **194**, 713-724.

Smith-Gill, S.J., Lavoie, T.B. and Mainhart, C. (1984). Antigenic regions defined by monoclonal antibodies correspond to structural domains of avian lysozyme. *J. of Immunol.*, **133**, 384-393.

Soderlind, E., Simonsson, A.C. and Borrebaeck, C.A.K. (1992). Phage display technology in antibody engineering: design of phagemid vectors and *in vitro* maturation systems. *Immun. Rev.* **130**; 109-123.

Somerville, J.E., Goshorn, S.C., Fell, H.P. and Darveau (1994). Bacterial aspects associated with the expression of a single-chain antibody fragment in *Escherichia coli. Appl. Microbiol. Biotechnol.* **42**; 595-603.

Spada, S., Krebber, C. and Pluckthun, A. (1997). Selectively infective phages. *Biol. Chemistry.*, **378**, No.6, 445-456.

Stanfield, B.B.and Jacobowitz, D.M., (1990). Antibody to a soluble-protein purified from brain selectively labels layer v corticofugal projection neurons in rat neocortex. *Brain research*, **531**, No.1-2, 219-224.

Suggs, S.V., Wallace, R.B., Hiros, E.T., Kawashima, E.H. and Itakura, K. (1981) Use of synthetic oligonucleotides as hybridisation probes. Isolation of cloned DNA sequences for human beta-2- microglobulin.*Proc. Natl. Acad. Sci.*, **78**, 6613-6617.

Sutor, D., (1958a). The structures of the pyrimidines: VI the structure of theophyllin. *Acta Crystallograpr.*, **11**, 83-87

Sutor, D., (1958b). The structures of the pyrimidines: VII the structure of theophyllin. *Acta Crystallograpr.*, **11**, 453-458

Swamy, K. H. S and Goldberg, A. L. (1981). *E. coli* contains eight soluble proteolytic activities, one being ATP dependent. *Nature (London),* **292,** 652-654.

Swamy, K. H. S and Goldberg, A. L. (1982). Subcellular distribution of various proteases in *Escherichia coli. J. Bacteriol.* **149,** 1027-1033.

Takkinen, K., Laukkanen, M.L., Sizmann, D., Alfthan, K., Immonen, T., Vanne, L., Kaartinen, M., Knowles, J.K.S, Teeri, T.T. (1991). An active single-chain antibody containing a cellulase linker domain is secreted by *Escherichia coli. Prot. Eng.*, **4**, No.7, 837-841.

Tavladoraki, P., Benvenuto, E., Trinca, S., De Martinis, D., Cattaneo, A. and Galeffi, P. (1993). Transgenic plants expressing a functional scFv antibody are specially protected from virus attack. *Nature* **366**; 469-472.

Thein, S.L., and Wallace, R.B. (1986). The use of synthetic oligonucleotides as specific hybridisation probes in the diagnosis of genetic disorders. *Human Genetic Diseases*: A practical approach., IRL Press.

Tomlinson, I.M., Walter, G., Marks, J.D., Llewelyn, M.B. and Winter, G. (1992) The repertoire of human germline VH sequences reveals about fifty groups of VH segmants with different hypervariable loops. *J. Mol. Biol.* **227**; 776-798.

Tormo, J., Blaas, D., Parry, N.R., Rowlands, D., Stuart, D. and Fita, I. (1994) Crystal structure of a human rhinovirus neutralizing antibody complexed with a peptide derived from viral capsid protein VP2. *EMBO J.* **13**; 2247-2256.

Tsumoto, K., Ogasahara, K., Ueda, Y., Watanabe, K., Yutani, K. and Kumagai, I. (1995) Role of Tyr residues in the contact region of antilysozyme monoclonal antibody HyHEL10 for antigen binding. *J. Biol. Chem.* **270** No. 31; 18551-18557.

Tsumoto, K., Nakaoki, Y., Ueda, Y., Ogasahara, K., Yutani, K., Watanabe, K. and Kumagai, I. (1994). Effect of the order of antibody variable regions on the expression of the single-chain HYHEL10 Fv fragment in *E. coli* and the thermodynamic analysis of its antigen-binding properties. *Biochemical and Biophysical Research Communications*, **201** No. 2; 546-551.

Tulip, W.R., Varghese, J.N, Webster, R.G., Laver, W.G. and Colman, P.M., (1992). Crystal-structures of 2 mutant neuraminidase antibody complexes with amino-acid substitutions in the interface. *J. Mol. Biol.*, **227**, No.1, pp.149-159

Ulrich, H.D., Patten, P.A., Yang, P.L., Romesberg, F.E. and Schultz, P.G. (1995). Expression stadies of catalytic antibodies. *Proc. Natl. Acad. Sci. USA* **92**; 11907-11911.

Vieira, J. and Messing, J. (1987). Production of single-stranded plasmid DNA. *Methods in Enzymology*, **153**, 3-11.

Verhoeyen, M., Milstein, C., and Winter, G. (1988). Reshaping human antibodies:grafting an anti-lysozyme activity. *Science* **239**; 1534-1536.

Wall, J.G. and Pluckthun, A. (1995). Effects of overexpressiong folding modulators on the *in vivo* folding of heterologous proteins in *Escherichia coli. Curr. Opin. Biotechnol.* **6,** 507-516.

Wall,J G and Pluckthun, A. (199). Different equilibrium stability behavior of ScFv fragments: Identification, classification, and improvement by protein engineering Protein Eng. **12**, 605-611

Walls, P.H. and Sterneberg, M.J.E. (1992). New algorithm to model protein-protein recognition based on surface complementarity. Applications to antibody-antigen docking. *J.Mol.Biol.*, **228**, 277-297.

Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T. and Winter, G. (1989) Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli. Nature* **341**; 544-563.

Waterhouse, P., Griffiths, A.D., Johnson, K.S., and Winter, G. (1993). Combinatorial infection and *in vivo* recombination: a strategy for making large antibody repertoires. *Nuc. Acid Res.* **21** No.9; 2265-2266

Webster,D.M., Henry, A.H. and Rees, A.R. (1994) Antibody-antigen interactions. *Cur. Op. Str. Biol.* **4**;123-129.

Whitney, G.K., Glick, B.R. and Robinson, C.W. (1989). Induction of T4 DNA-ligase in a recombinant strain of *Escherichia coli. Biotechn. and Bioeng.*, **33**, No.8, 991-998.

Wilson, I.A. and Stanfield, R.L. (1994). Antibody-antigen interactions: new structures and new conformational changes. *Cur. Op. Str. Biol.* **4**; 857-867.

Wilson, I.A. and Stanfield, R.L (1993). Antibody-antigen interactions. *Cur. Op. Str. Biol.* **3**;113-118.

Winter, G., Griffiths, A.D., Hawkins, R.E. and Hoogenboom, H.R. (1994). Making antibodies by phage display technology. *An. Review of Immunol.*, **12**, 433-455.

Winter, G. and Milstein C. (1991) Man-made antibodies. *Nature*, **349**, 293-299.

Witkowski, A., Daunert, S., Kindy, M. and Bachas, L. (1993). Enzyme linked immunosorbent assay for an octapeptide based on a genetically engineered fusion protein. *Anal. Chem,* **65**, 1147-1151.

Worn, A and Pluckthun,A. (1998). Mutual stabilization of V-L and V-H in single-chain antibody fragments, investigated with mutants engineered for stability Biochemistry 37, 13120-13127

Worn, A. and Pluckthun 1999 Different equilibrium stability behavior of ScFv fragments: Identification, classification, and improvement by protein engineering

, Biochemistry 38, 8739-8750

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains-nucleotide sequences of the m13MP18 and pUC19 vectors. *Gene*, **33**, 103-119.

Appendix I

A. Automated sequencing

The Perkin Elmer Applied Biosystems (ABI) 373A Automated Sequencer was used for all the sequencing reactions described in this thesis. The termination at each of the four nucleotides is characterised by a different coloured dye and all four reactions can be run in a single lane, rather than four separate lanes. Using dy-labelled terminators (DyeDeoxyTerminators, ABI) all four reactions can be performed simultaneously in the same tube (according to the manufacturer's instructions).

B. Software packages for sequencing analysis

There are several software packages for the anlysis of DNA sequences. If using an automated sequencer, the analysis software should be compatible with its output, so that the sequences can be directly imported. The ABI 373A automated sequencer software runs on an Apple Macintosh. The ABI sequence editor SeqEd (the updated version is called Sequence Navigator) was used for sequence analysis. SeqEd allows multiple

chromatograms from the same of different runs to be displayed simultaneously. The sequences can be compared and the differences between them highlighted. This was

particularly useful for comparing the Gloop2 and GlaMor scFv nucleotide sequences and identifying the mutated bases. SeqEd can also reverse complement sequences and produce amino acid translations. After performing the above manipulations using SeqEd the Gloop2 and GlaMor scFv full length sequences were saved and are shown below.

C. The Gloop2 and GlaMor scFv full sequences

The Gloop2 and GlaMor scFv sequences were translated in amino acid sequences. The protein sequences were compared and the different amino acids were marked with asterisks. In the format below all four nucleotide and protein sequences are shown as well as the differences between them. Both scFv constructs start with a *pel*B leader sequence at the N'-end followed by the V_H, the linker, the V_L, the myc-tag and the 6xhistine tag at the C'-end.

Appendix 1

```
                                    10        20        30        40        50
1... 50 Gloop2      TCTAGAATGA AATACCTATT GCCTACGGCA GCCGCTGGAT TGTTATTGCT
1... 50             S  R  M   K  Y  L  L   P  T  A   A  A  G   L  L  L  L
1... 50             ---------- ---------- ---------- ---------- ----------
1... 50 GlaMor      TCTAGAATGA AATACCTATT GCCTACGGCA GCCGCTGGAT TGTTATTGCT
1... 50             S  R  M   K  Y  L  L   P  T  A   A  A  G   L  L  L  L

                                    60        70        80        90       100
51...100 Gloop2     AGCTGCCCAA CCAGCGATGG CCCAGGTGCA GCTGCAGCAG TCTGGAACTG
51...100            A  A  Q   P  A  M   A  Q  V  Q   L  Q  Q   S  G  T
51...100            ---------- ---------- ---------- ---------- ----------
51...100 GlaMor     AGCTGCCCAA CCAGCGATGG CCCAGGTGCA GCTGCAGCAG TCTGGAACTG
51...100            A  A  Q   P  A  M   A  Q  V  Q   L  Q  Q   S  G  T

                                   110       120       130       140       150
101...150 Gloop2    AGTTGGCGAG GCCTGGGGCT TCAGTGAGGC TGTCCTGCAA GGCTTCTGGA
101...150           E  L  A   R  P  G   A  S  V  R   L  S  C   K  A  S  G
101...150           ---------- ---------- ---------- ---------- ----------
101...150 GlaMor    AGTTGGCGAG GCCTGGGGCT TCAGTGAGGC TGTCCTGCAA GGCTTCTGGA
101...150           E  L  A   R  P  G   A  S  V  R   L  S  C   K  A  S  G

                                   160       170       180       190       200
151...200 Gloop2    TACACCTTCA CAACCTTTGG TATAACCTGG GTGAAGCAGA GAACTGGACA
151...200           Y  T  F   T  T  F   G  I  T  W   V  K  Q   R  T  G  Q
151...200           ---------- ---------- --*--*---- ---------- ----------
151...200 GlaMor    TATACCTTCA CAACCTTTGG TAGATTCTGG GTGAAGCAGA GAACGGGACA
151...200           Y  T  F   T  T  F   G  R  F  W   V  K  Q   R  T  G  Q

                                   210       220       230       240       250
201...250 Gloop2    GGGCCTTGAG TGGATTGGAG AAATTTTTCC TGGAAATAGT AAGACTTACT
201...250           G  L  E   W  I  G   E  I  F  P   G  N  S   K  T  Y
201...250           ---------- ---------- *--------- ---------- -*--------
201...250 GlaMor    GGGCCTTGAG TGGATTGGAC ATATTTTTCC TGGAAATAGT TGGACTTACT
201...250           G  L  E   W  I  G   H  I  F  P   G  N  S   W  T  Y

                                   260       270       280       290       300
251...300 Gloop2    ACGCTGAGAG GTTCAAGGGC AAGGCCACAC TGACCGCAGA CAAATCCTCC
251...300           Y  A  E   R  F  K   G  K  A  T   L  T  A   D  K  S  S
251...300           ---------- ---------- ---------- ---------- ----------
251...300 GlaMor    ACGCTGAGAG GTTCAAGGGC AAGGCCACAC TGACCGCAGA CAAATCCTCC
251...300           Y  A  E   R  F  K   G  K  A  T   L  T  A   D  K  S  S

                                   310       320       330       340       350
301...350 Gloop2    ACCACAGCCT ACATGCAGCT CAGCAGCCTG ACATCTGAGG ACTCTGCCGT
301...350           T  T  A   Y  M  Q   L  S  S  L   T  S  E   D  S  A  V
301...350           ---------- ---------- ---------- ---------- ----------
301...350 GlaMor    ACCACAGCCT ACATGCAGCT CAGCAGCCTG ACATCTGAGG ACTCCGCCGT
301...350           T  T  A   Y  M  Q   L  S  S  L   T  S  E   D  S  A  V

                                   360       370       380       390       400
351...400 Gloop2    CTATTTCTGT GCAAGAGAGA TCCGCTACTG GGGCCAAGGG ACCAAGGTCA
351...400           Y  F  C   A  R  E   I  R  Y  W   G  Q  G   T  K  V
351...400           ---------- -------*-- *--*------ ---------- ----------
351...400 GlaMor    CTATTTCTGT GCAAGACATG ACTACTACTG GGGCCAAGGC ACCAAGGTCA
351...400           Y  F  C   A  R  H   D  Y  Y  W   G  Q  G   T  K  V
```

```
                                410        420        430        440        450
401...450 Gloop2    CCGTCTCCTC AGGTGGAGGC GGTTCAGGCG GAAGTGGCTC TGGCGGTGGC
401...450           T  V  S  S   G  G  G   G  S  G   G  S  G  S   G  G  G
401...450           ---------- ---------- ---------- ---------- ----------
401...450 GlaMor    CCGTCTCCTC AGGTGGAGGC GGTTCAGGCG GAAGTGGCTC TGGCGGTGGC
401...450           T  V  S  S   G  G  G   G  S  G   G  S  G  S   G  G  G

                                460        470        480        490        500
451...500 Gloop2    GGATCGGACA TCGAGCTCAC CCAGTCTCCA TCCTCCTTAT CTGCCTCTCT
451...500            G  S  D   I  E  L   T  Q  S   P  S  S   L  S  A  S  L
451...500           ---------- ---------- ---------- ---------- ----------
451...500 GlaMor    GGATCGGACA TCGAGCTCAC CCAGTCTCCA TCCTCCTTAT CTGCCTCTCT
451...500            G  S  D   I  E  L   T  Q  S   P  S  S   L  S  A  S  L

                                510        520        530        540        550
501...550 Gloop2    GGGAGAAAGA GTCAGTCTCA CTTGTCGGGC AAGTCAAGAA ATTAGTGGTT
501...550             G  E  R   V  S  L   T  C  R   A  S  Q   E  I  S  G
501...550           ---------- ---------- ---------- ---------- ----------
501...550 GlaMor    GGGAGAAAGA GTCAGTCTCA CTTGTCGGGC AAGTCAAGAA ATTAGTGGTT
501...550             G  E  R   V  S  L   T  C  R   A  S  Q   E  I  S  G

                                560        570        580        590        600
551...600 Gloop2    ACTTAAGCTG GCTTCAGCAG AAACCAGATG GAACTATTAA ACGCCTGATC
551...600           Y  L  S  W   L  Q  Q   K  P  D   G  T  I   K  R  L  I
551...600           *--------- ---------- ---------- ---------- ----------
551...600 GlaMor    TCTTAAGCTG GCTTCAGCAG AAACCAGATG GAACTATTAA ACGCCTGATC
551...600           F  L  S  W   L  Q  Q   K  P  D   G  T  I   K  R  L  I

                                610        620        630        640        650
601...650 Gloop2    TACGCCGCAT CCACTTTAGA TTCTGGTGTC CCAAAAAGGT TCAGTGGCAG
601...650            Y  A  A   S  T  L   D  S  G   V  P  K   R  F  S  G  R
601...650           ---------- ---------- ---------- ---------- ----------
601...650 GlaMor    TACGCCGCAT CCACTTTAGA TTCTGGTGTC CCAAAAAGGT TCAGTGGCAG
601...650            Y  A  A   S  T  L   D  S  G   V  P  K   R  F  S  G  R

                                660        670        680        690        700
651...700 Gloop2    AAGGTCTGGG TCAGATTATT CACTCACCAT CAGCAGCCTT GAGTCTGAAG
651...700             R  S  G   S  D  Y   S  L  T   I  S  S   L  E  S  E
651...700           ---------- ---------- ---------- ---------- ----------
651...700 GlaMor    AAGGTCTGGG TCAGATTATT CACTCACCAT CAGCAGCCTT GAGTCTGAAG
651...700             R  S  G   S  D  Y   S  L  T   I  S  S   L  E  S  E

                                710        720        730        740        750
701...750 Gloop2    ATTTTGCAGA CTATTATTGT CTACAATATC TTAGTTATCC GCTCACGTTC
701...750           D  F  A  D   Y  Y  C   L  Q  Y   L  S  Y   P  L  T  F
701...750           ---------- ---------- -*------*-- ------*--- --*-------
701...750 GlaMor    ATTTTGCAGA CTATTATTGT TGCCAAGATC TTAGTTTTCC GGACACGTTC
701...750           D  F  A  D   Y  Y  C   C  Q  D   L  S  F   P  D  T  F

                                760        770        780        790        800
751...800 Gloop2    GGTGCTGGGA CCAAGCTCGA GCTGAAACGG GCGGCCGCAG AACAAAAACT
751...800            G  A  G   T  K  L   E  L  K   R  A  A   A  E  Q  K  L
751...800           ---------- ---------- ---------- ---------- ----------
751...800 GlaMor    GGTGCTGGGA CCAAGCTCGA GCTGAAACGG GCGGCCGCAG AACAAAAACT
751...800            G  A  G   T  K  L   E  L  K   R  A  A   A  E  Q  K  L
```

```
                                 810        820        830        840        850
801...850 Gloop2    CATCTCAGAA GAGGATCTGA ATGGGGCCGC ACATCACCAT CATCACCATT
801...850             I  S  E    E  D  L    N  G  A    A    H  H  H    H  H  H
801...850           ---------- ---------- ---------- ---------- ----------
801...850 GlaMor    CATCTCAGAA GAGGATCTGA ATGGGGCCGC ACATCACCAT CATCACCATT
801...850             I  S  E    E  D  L    N  G  A    A    H  H  H    H  H  H

                                 860        870        880        890        900
851...855 Gloop2    AATAA
851...855           *   *
851...855           -----
851...855 GlaMor    AATAA
851...855           *   *
```

Appendix II

A. Immobilisation, binding and regeneration on the BIAcore

SPR analysis depends on immobilising (convalent binding) sufficient amounts of the ligand to the carboxymethyl dextran hydrogel. In the case of protein ligands, proteins usually possess sufficient surface lysine residues so that coupling to the carboxyl groups of the hydrogel can be accomplished without completely destroying the

ligand's antibody binding activity. Johnson *et al.* (1991) have shown that maximum protein binding to the hydrogel occurs at a pH below the protein's isoelectric point and at low ionic strength (10mM or less) where ionic interactions between the protein and hydrogel are favoured. Amine coupling of protein to EDC/NHS activated carboxyl groups of the hydrogel, the standard BIAcore activation chemistry, occurs best at the pH where the matrix amount of protein electrostatically interacts with the hydrogel. For most proteins, even those of low pI, the optimum pH is from 4.0-4.5, the optimum pH for EDC activation of carboxyl groups. The entire process is monitored by a sensorgram and the net target value of net covalent binding to the hydrogel should be from a minimum of

1000 to a maximum of 10000 RUs. At the lower end, there may not be sufficient range (in RUs) to cover the binding kinetics of 4-5 times concentrations of ligate. At the upper end, too much ligand (immobilised part) may be present on the biosensor, causing problems with mass transport at low analyte (injected part) concentrations.

The amount of ligand to be immobilised for a given application also depends on the relative sizes of ligand and analyte. For a large ligand, a given level in RU will represent a smaller number of analyte binding sites giving a correspondingly smaller maximum binding capacity in RU. Similarly, a large analyte will give a larger response than a small analyte for the same number of binding sites.

The apparent stoichiometry of the surface complex may be calculated from the saturating binding capacity of the surface if the molecular weights of analyte and ligand are known:

$$Stoichiometry = \frac{analyte\ response}{ligand\ responsel} x \frac{ligand\ MW}{analyte\ MW}$$

Ethanolamine is added after the ligand to deactivate any remaining active esthers.
Regeneration solution may be added to condition the surface before analyte interaction analysis. Regeneration refers to removal of non-covalenlty bound analyte: the covalently immobilized ligand cannot normally be removed without destroying the sensor chip surface. For antibody-antigen interaction, the surface can usually be regenerated by lowering the pH.

The analyte, introduced to the surface with the immobilised ligand, can be analyzed in crude sample, such as serum and cell culture supernatants. However, it should be free from free particles. Sample concentration and the required amount of substance varies between applications. Typical concentration of injected analyte is 2-100μg/ml.

B. Measuring association and dissociation rate constants.

When the analyte is injected in a discrete pulse across a ligand surface, the resulting sensorgram can be divided into three essential phases.

1. Association of analyte with ligand during sample injection
2. Equilibrium or steady state during sample injection, where the rate of analyte binding is balanced by dissociation from the complex.
3. Dissociation of analyte from the surface during buffer flow.

The three phases are illustrated in Fig. II.2.

The association and dissociation phase provides information on the *kinetics* of the analyte-ligand interaction (i.e. the rates of complex formation and dissociation). The equilibrium phase provides information on the *affinity* of the analyte-ligand interaction. The basic kinetic measurements are shown in Table II.1.

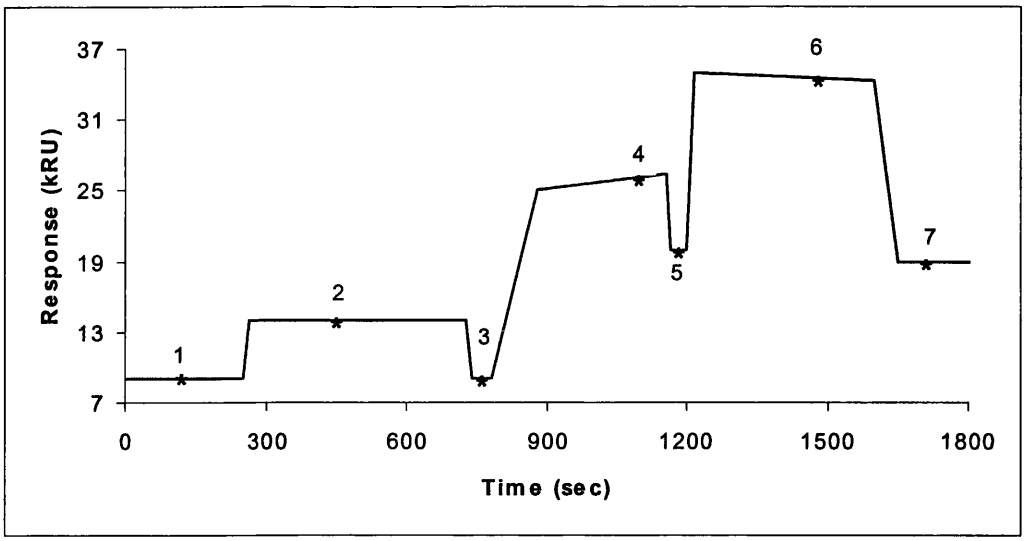


Fig. II.1: Typical sensorgram from immobilisation of monoclonal antibodies to sensor chip CM5 using standard amine coupling chemistry.

1. Baseline for unmodified sensor chip surface with continuous flow buffer.
2. Injection of NHS/EDC to activate the surface gives an increase in the SPR signal due to the change of bulk refractive index.
3. Baseline after activation. Activation of the surface has itself only a very slight effect on the SPR signal.
4. Injection of ligand (IgG) leads to electrostatic attraction and coupling to the surface matrix. At this point the ligand solution is still in contact with the sensor chip surface and the response includes both immobilised and non-covalently bound ligand.
5. Immobilised ligand before deactivation. The ligand solution has passed the sensor surface, and most of the protein that is not covalently bound is eluted.
6. Deactivation of unreacted NHS-esthers using 1M ethanolamine hydrochloride, adjusted to pH 8.5 with NaOH. The increased SPR signal is due to change in the bulk refractive index. The deactivation process also removes any remaining electrostatically bound ligand.
7. Immobilised ligand after deactivation.

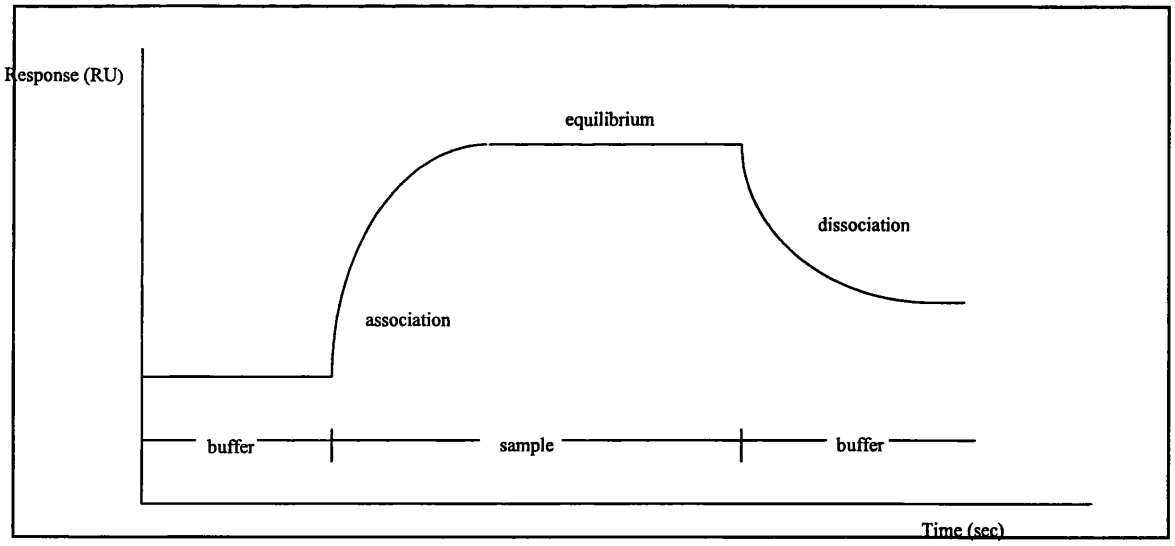


Fig. II.2: Schematic sensorgram, showing association, equilibrium and dissociation phases.

Basic Kinetic Measurements

A homogeneous 1:1 interaction on the sensor chip surface may be described by the equation:

$$A + B \underset{Koff}{\overset{Kon}{\rightleftharpoons}} AB$$

Where A is the analyte, B is the surface-bound ligand, K_{on} the association constant and K_{off} the dissociation constant.

Assuming first order interaction kinetics, the rate of association phase during sample injection is given by:

$$d[AB]/dt = K_{on}[A][B] - K_{off}[AB]$$

which may be expressed in terms of the SPR signal as:

$$dR/dt = K_a C Rmax - (K_{on} C + K_{off}) R$$

Respectively the dissociation phase is described by:

$$dR/dt = -K_d R$$

Where dR/dt is the rate of change of the SPR signal

C is the concentration of the analyte

Rmax is the maximum analyte binding capacity in RU

and R is the SPR signal in RU at time t.

Table II.1: The basic kinetic measurements.